G.E.C. Journal

VOL. XI. AUGUST, 1941 No. 4.

CONTENTS

PAGE

FRONTISPIECE: 40,000 kVA TURBO-ALTERNATOR SET IN A MUNICIPAL POWER STATION - - - - - - 222

G.E.C. HEAVY ALLOY - - - - - - - - 223
By G. H. S. Price, S. V. Williams and C. J. O. Garrard

ECONOMIC FACTORS OF POWER PRODUCTION BY STEAM PLANT 238
By B. Pochobradsky

BRIGHTNESS AND CONTRAST IN ILLUMINATING ENGINEERING 246
By R. G. Hopkinson, W. R. Stevens and J. M. Waldram

PHOSPHORESCENCE AND ITS APPLICATIONS - - - - 256
By J. N. Bowtell and E. E. Miles

TELEVISION FILM TRANSMITTERS USING APERTURED SCANNING DISCS, PART I. - - - - - - - 266
By D. C. Espley and D. O. Walter

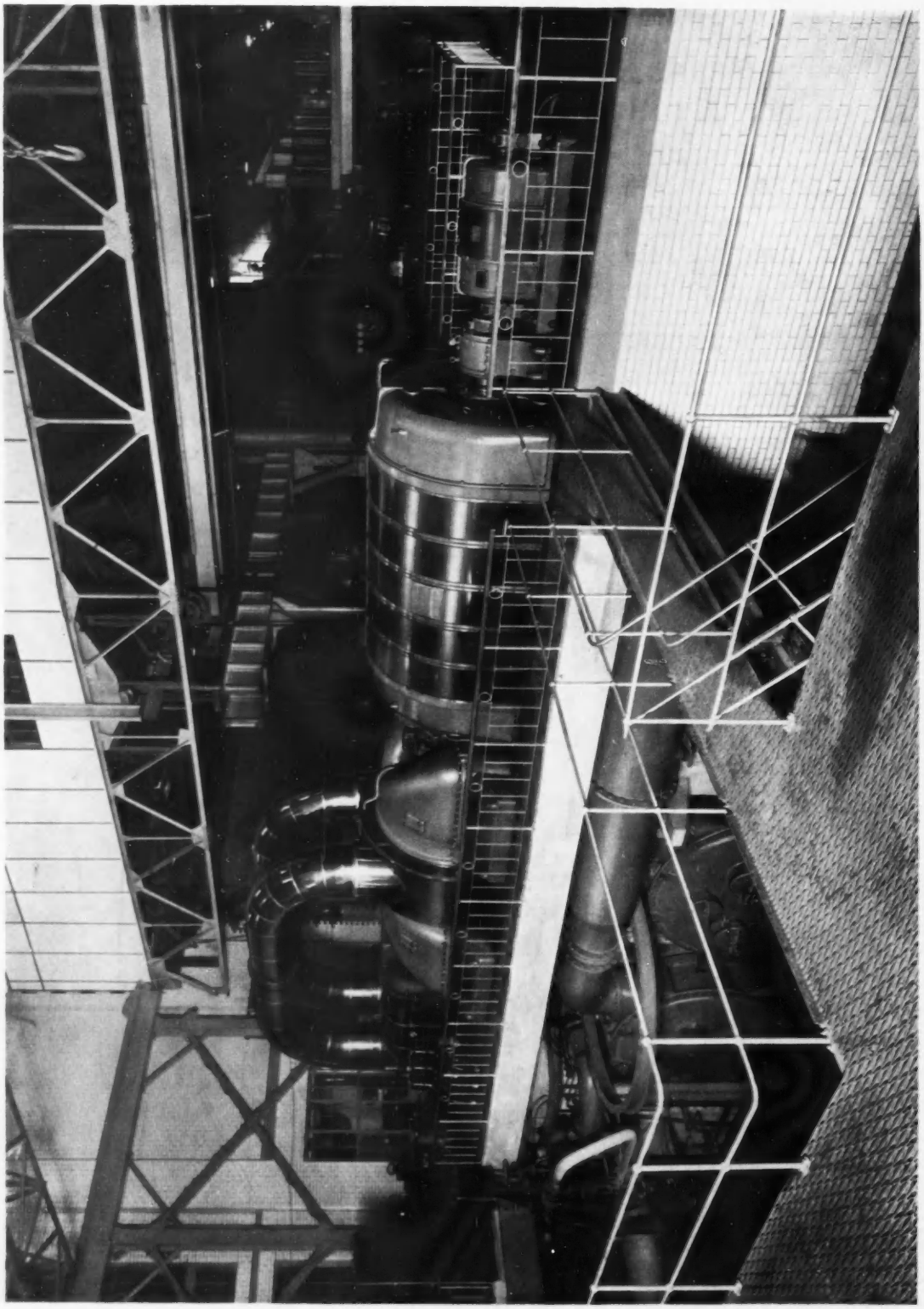
Published by

THE GENERAL ELECTRIC Co., LTD.
Magnet House, Kingsway, London, England.

COPYRIGHT.

PRICE 1/-

Copies can be obtained from the above address.



40,000 kVA TURBO-ALTERNATOR SET IN A MUNICIPAL POWER STATION.
A 40,000 kVA, 6,600 volt G.E.C.-Fraser & Chalmers turbo-alternator set recently put into service at a Municipal Generating Station. The turbine is supplied with steam at a pressure of 400 lbs. per sq. in., and drives the generator at 3,000 r.p.m.
In addition to the power unit, the equipment provided includes the main A.C. switchgear and the motors for the condenser auxiliaries.

G.E.C. Journal

VOL. XI. AUGUST, 1941. No. 4.

G.E.C. Heavy Alloy.

ITS PRODUCTION, PROPERTIES AND USES.

By G. H. S. PRICE, B.Sc. and S. V. WILLIAMS, B.Sc.,
G.E.C. Research Laboratories;
and C. J. O. GARRARD, M.Sc., A.M.I.E.E.,
Witton Engineering Works.

In this article, the authors describe the preparation, properties and industrial uses of an alloy which is fifty per cent heavier than lead and has the tensile strength of a good quality steel.

The alloy, which consists essentially of tungsten, is made by a powder metallurgical process which is fully described. Its high density makes it suitable for many screening purposes, such as the construction of radium containers, while its good resistance to arcing, conferred on it by its high tungsten content, makes it an ideal material for the arcing contacts of oil and air circuit breakers. In addition to these specialised uses, Heavy Alloy has found many applications as a general engineering material and some of these, together with an indication of future developments, are discussed.

INTRODUCTION.

SOME of the most outstanding metallurgical developments of recent years have been in connection with the production of light alloys which have played such an important part in the development of modern aircraft. These alloys are principally of aluminium and magnesium and their properties have virtually made them a series of entirely new structural materials. For certain purposes however, a *heavy* metal or alloy is required, and it is interesting to record the successful development of an alloy at the other end of the density scale, namely, G.E.C. Heavy Alloy.

Table I gives the densities of a number of metals and alloys, and from this it will be seen that of the common metals, lead is the only one which is substantially heavier than the general engineering materials. Lead, with its excellent corrosion resisting properties, is a very useful metal, but it is weak mechanically, and for this reason its use is limited to those applications where the stresses are low. The precious metals, such as gold and platinum, are heavy metals, being 2—3 times as heavy as iron or copper, but their high price prohibits their use for general engineering purposes. Tungsten however, has a density equal to that of gold and is also relatively cheap. A considerable tonnage of tungsten is used annually as an alloying element in the production of special steels, and for this purpose the metal is added in the form of powder or as a ferro-tungsten alloy. Tungsten is characterised by a very high melting point (3400 degs. C.) and this precludes the production of the pure metal by ordinary metallurgical processes.

Pure tungsten in the form of ductile wire or sheet, such as is used in the manufacture of electric lamps and radio valves, is produced by a powder metallurgical process. The successful development of the process about 30 years ago was one of the first and most important examples of this method of producing metals and alloys, a method which has been extended in many directions, particularly during the last few years. So far as tungsten is concerned, a bar of pressed powder is sintered to a high temperature and subsequently worked by swaging, drawing or rolling to produce the metal in wire or sheet form. The sintering temperature required is in the region of 3000 degs. C. and this is obtained by passing an electric current (about 2000 amps. per 1/16 square inch) through the bar. This method of

sintering however, limits the size of bar that can be conveniently handled, and having regard to the fact that it is only after a considerable amount of work has been done on the sintered bar that the metal approaches its theoretical density, it is evident that

TABLE I.

Densities of some Metals and Alloys.

Elektron (Magnesium Alloy) ..	1.8 gms./c.c.
Aluminium	2.7 ,,
Steel	7.8 ,,
Brass	8.4 ,,
Copper	8.9 ,,
Nickel	8.9 ,,
Molybdenum	10.0 ,,
Lead	11.4 ,,
HEAVY ALLOY	16.5—17.0 ,,
Tungsten	19.3 ,,
Gold	19.3 ,,
Platinum	21.5 ,,

fully dense tungsten cannot be produced in massive form. Although the production of large pieces of pure tungsten is not a practical proposition, it is clear that an alloy consisting mainly of tungsten would still have a density considerably higher than lead and would be very useful for the purpose for which such a material was originally required, namely radium containers.

In the first instance attempts were made to produce an alloy of tungsten and lead. Tungsten powder was mixed and pressed with a sufficient quantity of lead powder to fill the interstices between the tungsten particles, when the mixture was heated. It was found, however, that tungsten and lead do not wet each other and homogeneous masses could not be produced. It was known that nickel and tungsten alloy very readily, and experiments were made using mixtures of tungsten powder with 5—10 per cent nickel. On pressing and sintering to a moderate temperature, alloys were obtained with densities of the order of 16·5—17 gms./c.c., which is more than twice that of steel (7·8) and about 50 per cent greater than lead.[1, 2, 3] It was then found that the addition of a certain amount of copper assisted production, and although the quantities of nickel and copper can be varied, the material, which is known as "Heavy Alloy," usually contains 90 per cent tungsten, 7·5 per cent nickel and 2·5 per cent copper. In addition to its high density, Heavy Alloy has a tensile strength comparable with that of a good quality steel, and has consequently found applications other than that for which it was originally developed.

Various factors which influence the production and properties of Heavy Alloy are dealt with in the early part of this article, while the latter part is devoted to descriptions of some of the uses for which it has been successfully adopted.

GENERAL PREPARATION.

As has already been indicated, Heavy Alloy is made by a powder metallurgical process, and, although the various factors influencing the properties of the final material are dealt with in greater detail later, the general method is as follows:—The requisite quantities of tungsten, nickel and copper powders, of a suitable degree of fineness, are thoroughly mixed together in the dry state. A small quantity of a suitable binder, such as a solution of wax in benzol, is then added, and the mixing continued until all the solvent has evaporated. The binder makes it possible to press coherent masses at much lower pressures than would otherwise be necessary, and for general purposes a pressure of from 5 to 10 tons per sq. in. is quite satisfactory.

The pressed pieces, which are strong enough to be handled, are supported on steel plates or refractory blocks and slowly heated to a temperature of approximately 1000 degs. C. in a muffle furnace. A reducing

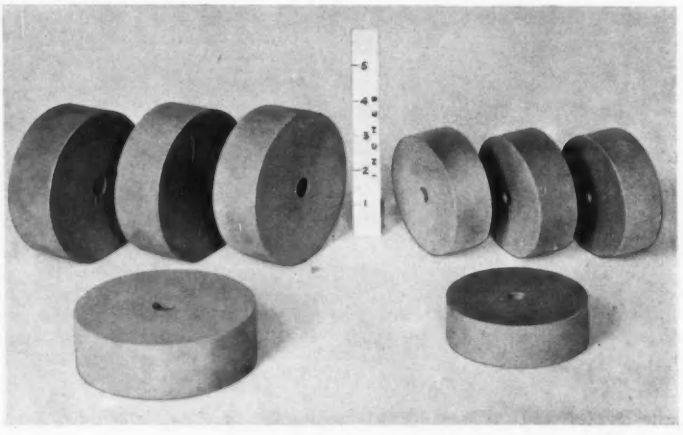


Fig. 1.—Heavy alloy rings before and after sintering.

atmosphere has to be maintained in the furnace and it is necessary to carry out the initial stages of this firing slowly if disruption of the pieces, due to too rapid evolution of the binder, is to be avoided. During this pre-firing treatment there is little tendency for the pressings to sinter and stick together and consequently it does not matter if they are in contact with one another in the furnace. The second stage of the firing, namely, heating the pressings to a temperature at which a liquid phase is formed, is much more critical and it is important that the pressings should not be in contact with one another during the high temperature firing. During this second firing, which must also be done in a reducing atmosphere, a linear shrinkage of up to 20 per cent occurs on all dimensions and this, of course, has to be allowed for in the designing of

the dies. Fig. 1, which illustrates the shrinkage, shows rings which were required about 3in. in diameter, but before firing had to be nearly 4in. diameter. When complicated shapes are required

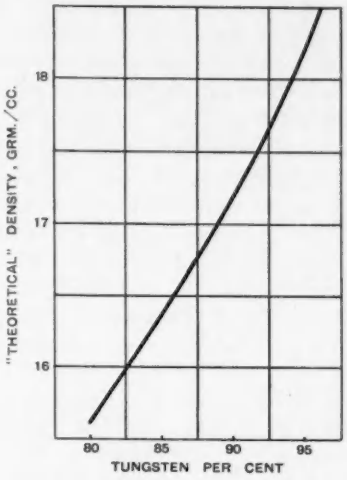


Fig. 2.—Theoretical densities of copper-nickel-tungsten alloys.

which cannot conveniently be pressed in one operation or when the quantity is not sufficient to justify the cost of a special die, it is usual to do a certain amount—if not all—of the necessary shaping after the prefiring treatment. The material at this stage is brittle and has to be handled with care although it is sufficiently strong to be chucked in a lathe or cut with a hacksaw. Heavy Alloy prepared by the process which has been briefly outlined is entirely free from porosity and has the properties of alloys prepared by the more conventional methods of melting and casting.

Effect of Composition.

The changes in density and microstructure which take place during sintering have been followed for a number of alloys of different composition. Nickel alloys very readily with tungsten, and it was found that a mixture of tungsten powder with approximately 10 per cent of nickel powder when pressed and then sintered at about 1500 degs. C. gave a satisfactory alloy having a density of over 17 gms./c.c. The addition of copper to nickel brings about a progressive lowering in its melting point, and it seemed reasonable to suppose that if part of the nickel were replaced by copper, it would be possible to obtain a satisfactory alloy at a lower sintering temperature. To a large extent, this proved to be true, and in addition to the lower sintering temperature, a marked improvement was obtained in the properties of the alloy.

In the case of a simple mixture consisting of x per cent of a metal of density d_1 and y per cent of a metal of density d_2 and z per cent of a metal of density d_3, the "theoretical" density d_t of the mixture is given by :—

$$d_t = \frac{100}{\frac{x}{d_1} + \frac{y}{d_2} + \frac{z}{d_3}}$$

The density of tungsten is 19·25 gms./c.c., while the densities of nickel and copper are both 8·9 gms./c.c. The "theoretical" densities of mixtures containing various amounts of tungsten are shown in fig. 2.

Since the lowest value of density of practical interest is about 16 gms./c.c., only alloys containing more than 80 per cent tungsten have been investigated in detail. A series of alloys containing different percentages of tungsten, nickel and copper was made under identical conditions, and the resultant density measured in each case. The results are shown in fig. 3. The numerals in the circles indicate the measured densities corresponding to the different compositions. Curves have been drawn to show compositions of approximately equal density, and the lines marked at 45 degs. show the "theoretical" densities for the various mixes. From an examination of fig. 3, certain fundamental conclusions regarding the effect of composition on the density of the alloys can be drawn, although it must be emphasised that the figures given apply only to the

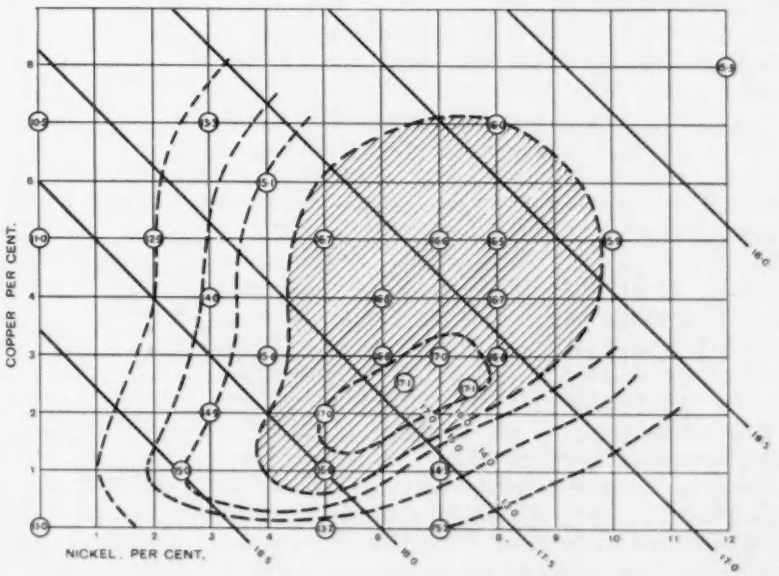


Fig. 3.—Effect of composition on density of copper-nickel-tungsten alloys.

particular conditions of the test. Different sintering conditions as regards time and temperature would give a slightly different, but similar, diagram. Certain conclusions, however, may be drawn which are independent of time and temperature of sintering, and these may be summarised as follows :—

(i). For a given percentage of tungsten, the maximum density is obtained when the ratio of nickel to copper is about 3 : 1.

(ii). Alloys in which the percentage copper exceeds that of nickel, do not generally sinter satisfactorily, and are liable to be porous.

(iii). The alloys approach their theoretical densities more closely as the percentage of tungsten is reduced, but the theoretical values, at the same time, become less.

Microexamination of the polished sections of these alloys has thrown considerable light on the mechanism of sintering and a theory has been put forward which satisfactorily accounts for all the observed facts. This is discussed in detail later.

The alloys containing tungsten and copper only do not show any appreciable shrinkage or change in density during sintering. The only effect of the sintering is that the molten copper "wets" the tungsten particles and cements them together. Alloys containing tungsten and nickel only, show a small shrinkage, and after polishing they have a more metallic appearance than those containing copper only. The nickel segregates into comparatively large areas, so that the tungsten particles are not completely wetted by it, and a large amount of porosity is apparent. This is doubtless due to the melting point of the nickel phase being raised above the sintering temperature by the solution of tungsten in nickel, so that eventually no liquid phase is present at the sintering temperature. As already mentioned (p. 224) satisfactory alloys *may* be obtained with tungsten and nickel only, but a higher temperature is required.

In alloys containing both nickel and copper, two factors are of importance : the ratio of nickel to copper and the total percentage of these metals present. The optimum nickel-copper ratio has already been given as about 3 to 1 and for the particular sintering conditions used for the series recorded in fig. 3, the optimum tungsten content is between 89 and 93 per cent. If it is borne in mind that a liquid phase is present during sintering, it will be realised that the *amount* of liquid must not be very great or there will be a danger of the whole body becoming plastic and losing its shape. This is often found to be the case when the combined nickel and copper additions exceed 10 per cent. On the other hand the sintering mechanism does not proceed within a reasonable time or at a practicable temperature if the nickel plus copper content is less than 5 or 6 per cent.

The optimum value for the nickel-copper ratio is determined by two factors, both of which are related to the copper content. Whilst the addition of copper lowers the melting point of the nickel and thus ensures the presence of a liquid phase which is able to penetrate between the tungsten grains, it, at the same time, reduces the solubility of tungsten in this phase, which, as will be shown later, is an essential part of the sintering process. Too high a proportion of copper is, therefore, a disadvantage. From a consideration of these facts it is possible to determine the most suitable composition from a practical point of view and the composition which has now been standardised for Heavy Alloy production is 90 per cent tungsten, 7·5 per cent nickel, 2·5 per cent copper. Although a 75 per cent nickel, 25 per cent copper alloy melts at 1350 degs. C., this figure is raised by the addition of tungsten, as shown in fig. 4, so that a temperature in excess of 1350 degs. C. is necessary if satisfactory alloys are to be obtained. For practical purposes it is considered that Heavy Alloy should have a density of not less than 16 gms./c.c. and it will be observed that all compositions within the shaded area on fig. 3 satisfy this condition.

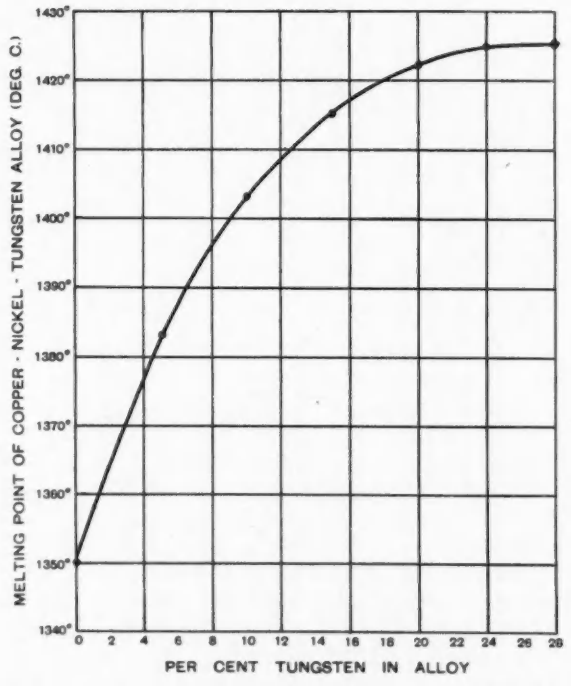


Fig. 4.—Effect of tungsten in solution on the melting point of a 3 : 1 Ni : Cu alloy.

Effect of Purity of Powders.

The density of any particular alloy is very dependent on the purity of the metal powders used. That this should be so is fairly obvious when it is considered that the presence of 2 per cent of an

impurity, with a density like that of aluminium or silicon, will reduce the theoretical density of an alloy with a nominal tungsten content of 90 per cent from 17·2 gms./c.c. to 15·6 gms./c.c. owing to the enormous difference in the volume of the impurity compared with that of the same weight of tungsten. This is an important point when considering the choice of raw materials for the production of Heavy Alloy and particularly the tungsten powder which, it is evident, should be of good quality. When considering a particular batch of tungsten it is, however, important that the composition of any impurities should be considered in addition to the actual quantity present. Thus a small percentage of tungstic oxide is not detrimental since this would be reduced to tungsten during the pre-firing stage.

Effect of Tungsten Particle Size.

It will be shown later that an essential factor in the elimination of voids during sintering is that the portion of the alloy which is liquid during sintering should not only wet the tungsten, but actually dissolve it. It would be expected, therefore, that the process would be more rapid the greater the percentage of fine particles, and this is confirmed by experiment. Three portions of a sample of tungstic oxide were reduced by heating in hydrogen, the time and temperature of reduction as well as the quantity of oxide treated being varied so as to produce a wide variation in the particle size of the resulting metal powders. Measurements of particle diameter gave the distributions recorded in Table II.

TABLE II.

TUNGSTEN POWDER	Particles within the limits per cent.			
	0—1 μ	1—2 μ	2—5 μ	>. 5 μ
Fine.. ..	98	2	0	0
Medium ..	13	54	33	0
Coarse ..	1	11	27	61

1 μ = 0.001 mm.

Alloys made up from each of these three grades of tungsten were sintered at different temperatures and the percentage shrinkages and resulting densities measured in each case.

From the results of this experiment it was evident that the finer the tungsten powder, the lower the sintering temperature necessary to produce an alloy of high density, *i.e.*, free from voids. Firing for ½ hour at the normal sintering temperature gave densities of 17·0, 16·9 and 15·6 gms./c.c. respectively for alloys made from fine, medium and coarse tungsten, while the corresponding shrinkages were 20·1, 16·6 and 14·5 per cent. When coarse tungsten was used, a density of 16·7 gms./c.c. and a shrinkage of 16·8 per cent were obtained only after prolonged sintering at a temperature much higher than that normally used. The actual particle size of the nickel and copper powders is not so important, provided they are of such a fineness that a uniform mixture with the tungsten powder is obtained.

Effect of Pressure.

Reference has already been made to the fact

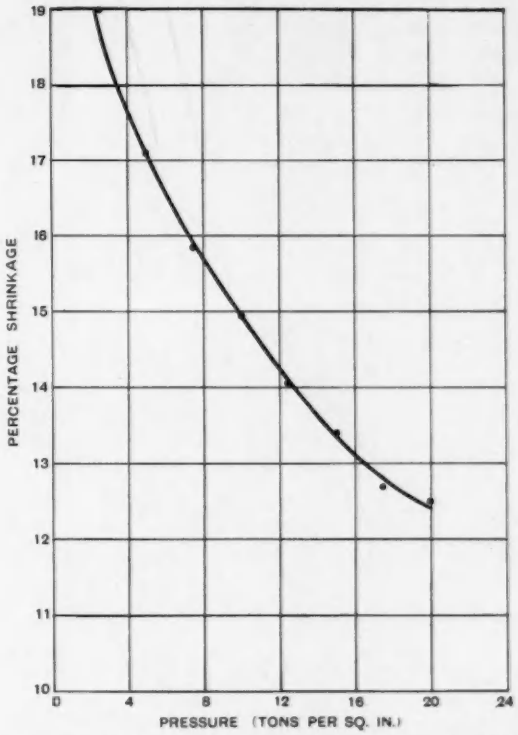


Fig. 5.—Effect of initial pressure on percentage shrinkage.

that a linear shrinkage of up to 20 per cent occurs on all dimensions during the sintering of Heavy Alloy. This shrinkage may, to a certain extent, be controlled by variations in the initial pressure. The higher the pressure the smaller will be the shrinkage, but the final density will not be appreciably affected. The effect of pressure is shown in fig. 5. For normal purposes a pressure of 5 tons/sq. in. is adopted, the percentage shrinkage in this case being about 17—17·5 per cent. There are certain practical limits to the depth of pressings, not only on account of the difference in volume between the unpressed and pressed powder, but because of the difficulty of transmitting a uniform pressure throughout any appreciable depth of powder. As shown in fig. 5 the percentage shrinkage is dependent on the initial pressure and consequently a pressing which is not uniformly pressed will not shrink uniformly. This fact is of importance when producing pieces to close tolerances in order to avoid subsequent machining and

it has to be taken into account in the design of dies. For simple shapes, such as cylinders, the difficulty may, in part, be overcome by means of a double acting type of press in which pressure can be applied from both ends, but even so, if distortion is to be avoided, the length of a cylinder should not be more than about twice its diameter.

Effect of Temperature of Sintering.

The shrinkage of Heavy Alloy does not progress uniformly with increasing temperature. There is a critical sintering temperature, which must be passed if satisfactory alloys are to be obtained, and this temperature corresponds with the appearance of a liquid phase and depends on the composition of the alloy. If an alloy of tungsten and nickel only is

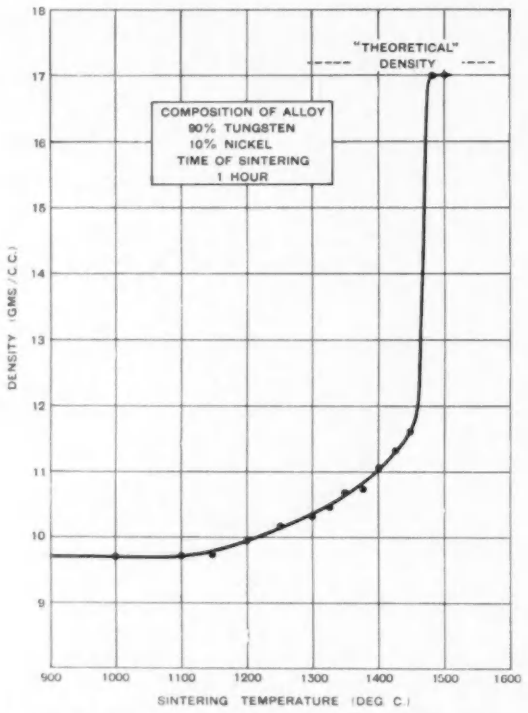


Fig. 6.—Effect of sintering temperature on density.

sintered, the fact that the nickel does not melt below 1450 degs. C. would preclude a dense body being obtained below this temperature. The fact, too, that molten nickel has been shown to dissolve tungsten[4, 5, 6] would indicate that a temperature appreciably higher than 1450 degs. C. is required if only tungsten and nickel are used. This is borne out by experiment, and the very rapid change in density as a result of sintering occurs between 1450 degs. C. and 1500 degs. C., as is shown in fig. 6. Very similar curves are obtained when copper is present, although the critical temperature is lower, since the presence of the copper lowers the melting point of the nickel. In order to investigate the microstructural changes which occur during sintering, a series of pressings, having the standard Heavy Alloy composition, was sintered at various temperatures. The resulting specimens were prepared for microexamination by polishing and then etching with boiling hydrogen peroxide, which etches tungsten but does not attack the nickel-copper phase. The microstructures of this series are shown in figs. 7—16.

Fig. 7 shows clearly the original components which were identified by their colour and have been marked. The subsequent photomicrographs show, first of all, the solid diffusion between the copper and nickel grains (fig. 8), then the cohesion between this nickel-copper phase and the tungsten (fig. 9) and finally the large increase in the size of the tungsten particles (figs. 10—12). Fig. 11 shows the typical microstructure of fully sintered Heavy Alloy, while fig. 12 shows a further change in structure which may be obtained by prolonged heating at the full sintering temperature. As the sintering proceeds and the density increases, the porosity will, of course, be reduced and this is indicated by the gradual disappearance of the black areas (which are voids) in figs. 13—16.

The compositions of the two phases now apparent were very easily determined. If the alloy be treated with aqua regia for about 24 hours, the whole of the nickel-copper phase can be dissolved and the rounded grains left unattacked as a residue. On analysis, these grains proved to be pure tungsten, and an X-ray diffraction pattern showed that each grain was a single tungsten crystal. The composition of the matrix was also determined from the X-ray diffraction pattern and this was found to be a nickel-copper alloy containing 23 per cent tungsten. The percentage of tungsten will be dependent on the relative proportions of nickel and copper in the alloy and will be lower, the higher the copper content. The structure of Heavy Alloy, therefore, consists of rounded single-crystal grains of tungsten, with diameters from 50 to 100 times those of the original tungsten grains, embedded in a matrix which consists of a nickel-copper alloy saturated with tungsten. There does not appear to be any advantage in prolonging the sintering to obtain the very large grain structure given in fig. 12.

The appearance of the liquid phase sets a lower limit to the sintering temperature and, although the upper limit is not so critical, there are a number of factors which make it desirable that the sintering should be carried out as near to the critical temperature as possible. For example, there is the danger of disruption of the material if the temperature is too high, caused by impurities boiling out. In addition to this, the copper-nickel-tungsten phase becomes more fluid with increasing temperature and consequently there is a danger of the Heavy Alloy distorting.

Effect of Time of Sintering.

Any consideration as to the effect of time of sintering must necessarily be connected with the temperature employed. It will be clear from the preceding paragraph that if the sintering temperature is not sufficiently high for the liquid phase to be present, then the sintering will never be complete however long the time. Once this critical temperature is reached sintering proceeds fairly rapidly, although with certain compositions a considerable time is necessary in order to attain the full density. Table III gives the results obtained with a 93 per cent tungsten alloy (theoretical density 17·8 gms./c.c.) sintered for various periods of time.

Prolonged sintering *below* the temperature at which the liquid phase appears rarely gives a density above 12 gms./c.c.

Not only is the time of sintering dependent on the temperature, but it is also very dependent on the composition of the alloy. Thus if the alloy contains a larger proportion of the nickel-copper phase, sintering may be completed very much more rapidly, but this saving in time is largely offset by the possibility of distortion due to over-sintering. Finally, reference has already been made (p. 227) to the fact that the finer the tungsten particles, the more rapid will be the sintering.

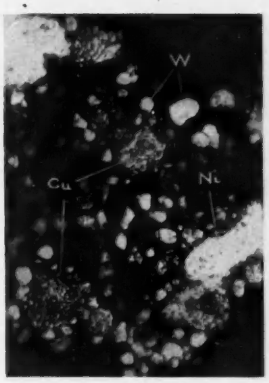


Fig. 7

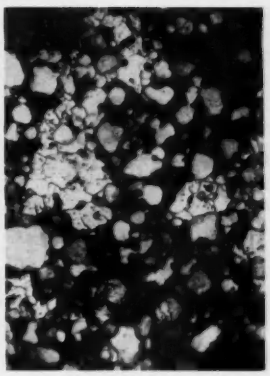


Fig. 8

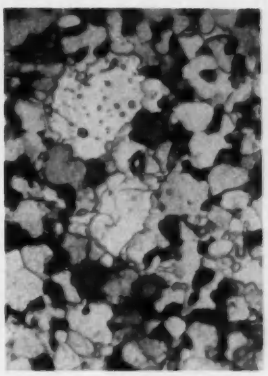


Fig. 9

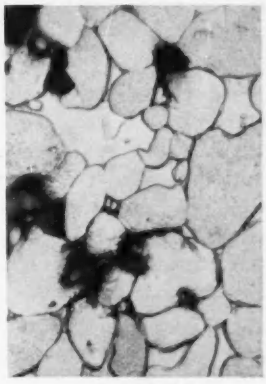


Fig. 10

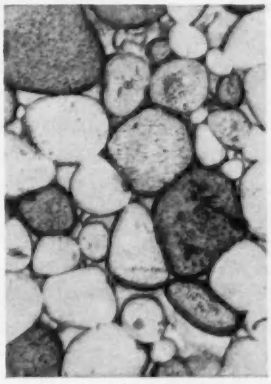


Fig. 11

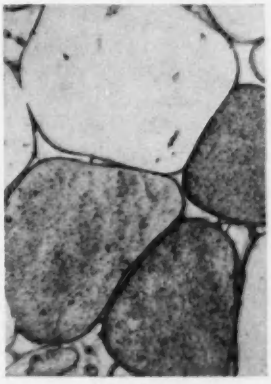


Fig. 12

Figs. 7, 8, 9, 10, 11 and 12.—Photomicrographs showing the effect of temperature on the progress of sintering of a tungsten-nickel-copper alloy (magnification × 300).

TABLE III.

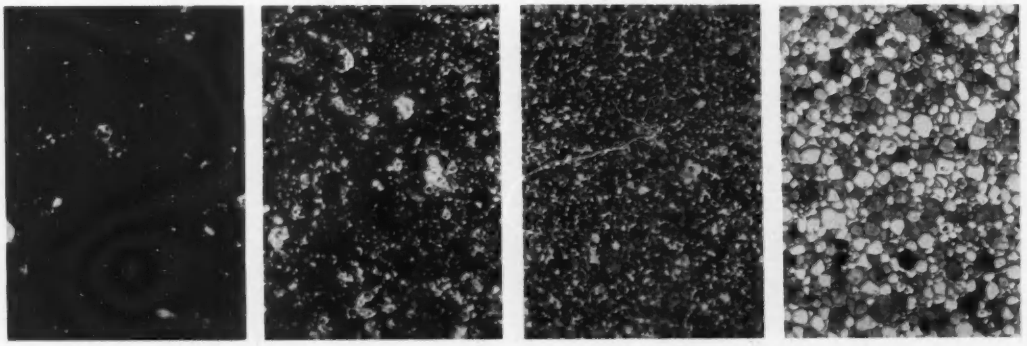
Effect of Sintering Time on Density.

Time of Sintering	Density—gms./c.c.
5 seconds	13.9
1 minute	14.1
5 minutes	15.0
¼ hour	16.3
½ hour	16.7
1 hour	17.2
6 hours	17.8

Effect of addition of Molybdenum.

Owing to the presence of the liquid phase during the sintering of Heavy Alloy, it is not practicable to reduce the quantity of tungsten to below 80 per cent. In the normal way, there is, of course, no object in doing this, since alloys with densities as high as possible are required and for an alloy containing 80 per cent tungsten, the "theoretical" density is only 15·6 gms./c.c. There are, however, cases where alloys of some definite density (say, above those of the common engineering materials and below

the standard Heavy Alloy range), are required and it is in this connection that the addition of molybdenum has proved useful. Molybdenum, like tungsten, is a highly refractory metal (m.p. 2450 degs. C.) but its density is appreciably less, being only 10·0 gms./c.c. as compared with 19·25 gms./c.c. for tungsten. Furthermore, molybdenum possesses all the properties necessary for the sintering mechanism to proceed, and it may thus be substituted for the tungsten in Heavy Alloy, the only effect being on the density of the finished product. The expression *"Heavy Alloy"* is only considered to be justifiable when the density of the alloy is greater than 15 gms./c.c. and for all practical purposes 16 gms./c.c. is taken as the lower limit. Excluding the use of lead, this leaves a big gap in the density scale between the common metals and alloys (even including the high tungsten steels) and the lower limit at which it is practicable to make Heavy Alloy. The substitution of molybdenum for a part, or all, of the tungsten in heavy alloy has made it possible to prepare a complete range of sintered alloys with densities ranging from about 10 up to 17 gms./c.c.



Figs. 13, 14, 15 and 16.—Photomicrographs showing the effect of temperature on the reduction in porosity during sintering (magnification × 60). The black portions are the porous areas.

THEORETICAL CONSIDERATIONS.

From the data given in the previous paragraphs, it appears that sintering of Heavy Alloy involves the following stages:

Between 1000 degs. and 1100 degs. C., the nickel and copper particles become alloyed by diffusion, and some cohesion with the tungsten grains takes place. Very little further change occurs until the temperature reaches the melting point of the nickel-copper alloy, but above this temperature a very remarkable change takes place. The tungsten particles are rapidly dissolved by the liquid phase, but tungsten is re-precipitated on certain nuclei, which develop into the characteristic large, rounded grains. This process continues until all the original fine tungsten particles have disappeared, and are replaced by fairly uniform grains with diameters about 100 times those of the original particles. At the same time all the remaining voids are eliminated.

That the tungsten grains grow by this process of solution and precipitation is confirmed by the behaviour of alloys containing copper only. Here there is no doubt of the presence of a liquid phase which wets the tungsten, but there is no solubility and no grain growth. With nickel only there is the necessary solubility, but, unless the temperature is above 1450 degs. C., there is no liquid phase and the process of growth is slow and incomplete. When the ratio of nickel to copper is less than 1 to 1, the solubility is so much reduced that the process is again very slow.

At first it is not obvious why the tungsten grains should dissolve, only to be thrown out of solution

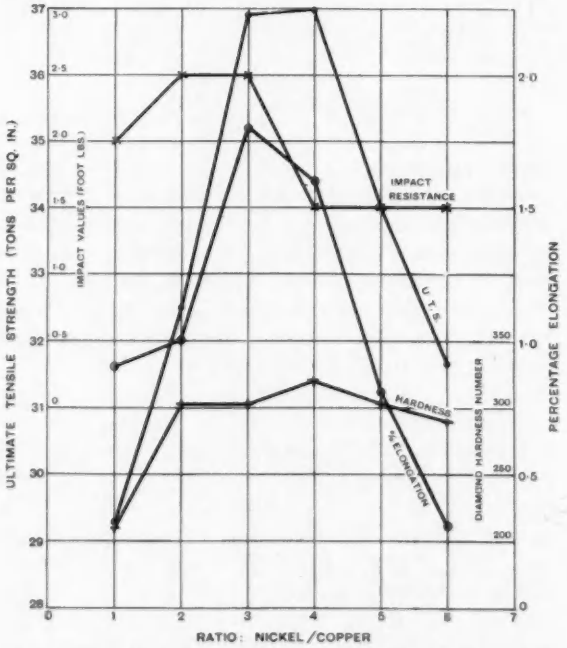


Fig. 17.—Effect of nickel-copper ratio on the mechanical properties of a 90 per cent tungsten alloy.

again. The explanation is to be found in the surface energy of the system. It has been well established that, for very small crystals, the interfacial tension between the crystals and their saturated solution is

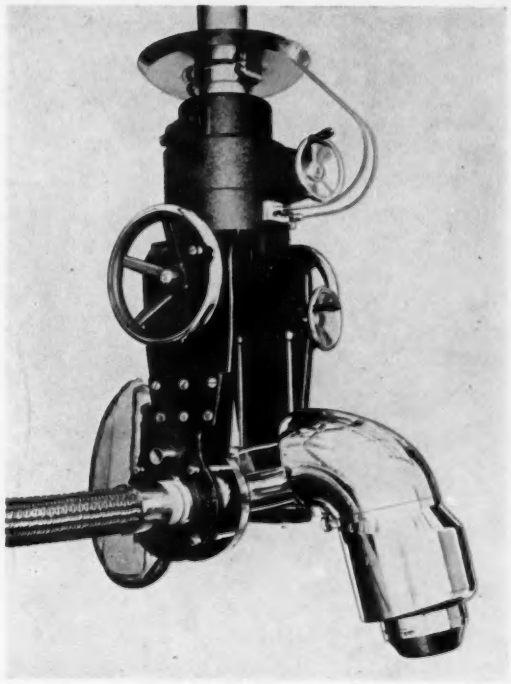


Fig. 18.—A 5 gm. radium unit with pneumatic transference of radium.

greater than that for large crystals.[7] This is the reason why small crystals are more soluble than large ones. The surface tension is related to the solubility by the equation[8]

$$\sigma = \frac{RT\,\Delta\,r}{2M} \log \frac{L_w}{L}$$

where σ is the interfacial tension, Δ is the density and M the molecular weight of the solid, L is the normal solubility, and L_w the solubility of small particles of radius r. Measurements made in different ways and with different soluble salts agree in demonstrating that this effect is only shown by crystals whose diameters are less than 1 μ, and that below this size it increases very rapidly as the diameter decreases. For example, the solubility in water of gypsum crystals 0·6 μ in diameter is 18 per cent greater than that of crystals 4 μ in diameter, which have the normal solubility.[9]

These facts offer a simple explanation of the changes observed in the alloys under investigation. The fine tungsten particles dissolve in the copper-nickel phase to give a solution which is supersaturated with respect to the larger tungsten grains, which therefore act as nuclei for precipitation and develop into the large, rounded grains. The observed effect of the size of the tungsten powder used is in agreement with this hypothesis.

If this picture is correct, we should expect that the type of structure shown by the copper-nickel-tungsten alloys would also be obtained with other alloy systems which fulfil the same equilibrium conditions, and would not be found among those

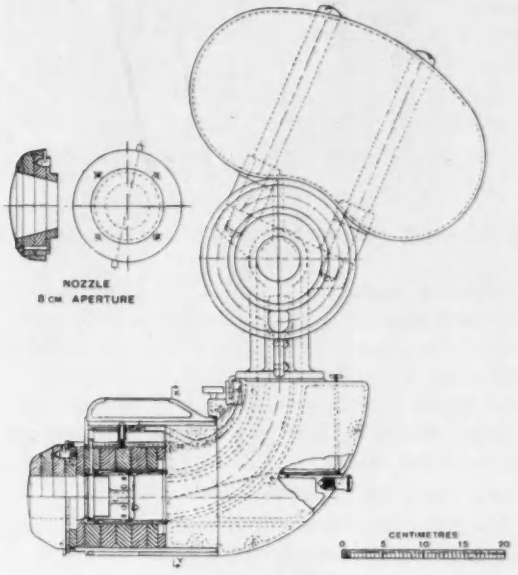


Fig. 19.—Section of radium unit showing use of Heavy Alloy in the container and nozzle. The shaded areas indicate the positions where Heavy Alloy is used.

which do not. This has been confirmed by experiment for a number of alloy systems, and it would appear that the only conditions which it is necessary to stipulate are that there must be present at the time of sintering, solid particles (of a certain fineness) and a liquid which is partially capable of dissolving them. It has been suggested by W. D. Jones[10] that it should be possible to bring about sintering conditions of this type in any alloy system simply by working between the solidus and the liquidus, provided that the liquid present is formed by some interaction between the solid particles and some other phase. This has definitely been confirmed for the following binary systems :—

(*a*) 90 per cent copper + 10 per cent tin, sintered at 925 degs. C.

(*b*) 80 per cent copper + 20 per cent silver, sintered at 900 degs. C.

(*c*) 80 per cent iron + 20 per cent copper, sintered at 1400 degs. C.

In each of these three cases, the typical Heavy Alloy structure of large rounded grains in a continuous matrix, was obtained. Other authors[11, 12, 13] have observed this type of structure, and in every case a liquid phase has been present. It is not, of course,

suggested that a liquid phase must be present in all cases of sintering when shrinkage occurs. Indeed, in the sintering of pure tungsten, where there is no liquid phase present, a shrinkage of the order of 17 per cent is obtained, but in this case the density of

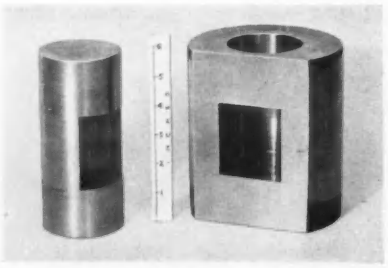


Fig. 20.—Machined components for a radium container body.

the sintered product is substantially lower than the theoretical density, and, in fact, it is only after a considerable amount of work has been done on the metal that fully dense tungsten is obtained. In other words the fact that shrinkage occurs during sintering is not in itself evidence of the complete elimination of voids.

PROPERTIES OF HEAVY ALLOY.

The more important physical properties of Heavy Alloy are summarised in Table IV.

TABLE IV.

Physical Properties of Heavy Alloy.

Tensile Strength—tons per sq. in.	40
Yield Point—tons per sq. in. ..	37
Elongation—per cent on 1 inch ..	4
Elastic Modulus—lb. per sq. in. ..	32×10^6
Brinell Hardness No.	250—290
Specific Gravity	16.3—17.0
Weight —lb. per cubic inch ..	0.6
Coefficient of expansion	5.6×10^{-6}
Thermal Conductivity—C.G.S. units	0.25
Specific Resistance—ohm cm^{-3}..	1.16×10^{-5}
Electrical Conductivity—ohm $^{-1}$ cm^{-1}	0.86×10^5

Thus the tensile strength and the yield point of Heavy Alloy are comparable with those of a high grade steel. The thermal conductivity is twice that of steel, but the coefficient of expansion is only about half. Heavy Alloy is very resistant to atmospheric corrosion. It may also be plated with any of the usual finishes such as nickel, cadmium or chromium.

The strength of Heavy Alloy, of any particular composition, is a function of the density and therefore increases as the sintering proceeds. The effect of sintering at the same temperature for various times on the density and strength of the resulting alloy, is shown in Table V.

TABLE V.

Effect of Sintering Time on Tensile Strength.

Time at Sintering Temperature	Density gms./c.c.	Ultimate Tensile Strength (tons/sq. in.)
½ minute	12.6	18.8
¼ hour	13.9	25.8
½ hour	15.0	29.2
¾ hour	15.6	33.4
1¼ hours	16.4	38.0
8 hours	16.8	41.5

In keeping with what has already been said regarding the necessity for sintering at a high enough temperature to ensure the presence of a liquid phase, it may be noted that prolonged sintering *below* this temperature yields a porous material whose strength does not exceed 25 tons/sq. in. The increase in the tensile strength from 38 to 41·5 tons/sq. in. as given in Table V, may be attributed to the final elimination of voids as indicated by an increase of 16·4 gms./c.c. to 16·8 gms./c.c. in the density. It is evident, thus, that the very large single crystals of tungsten which prolonged sintering produces (see fig. 12) do not exert much influence on the mechanical properties.

Little need be said regarding the effect of composition on the properties of Heavy Alloy since much of this will be obvious, following a consideration of the effect of composition on sintering and

TABLE VI.

Effect of Nickel-Copper ratio on Tensile Strength.

Tungsten Percentage	Nickel-Copper Ratio	Ultimate Tensile Strength tons/sq. in.	Density gms./c.c.
90	1	29.2	15.2
90	2	32.4	16.3
90	3	36.9	16.3
90	4	37.0	16.2
90	5	34.0	16.1
90	6	31.7	16.3

Fig. 21.—A group of Heavy Alloy contact tips for use on oil circuit breakers.

on the final product as given earlier. It has, however, been found that with any given tungsten content the maximum tensile strength is obtained when the nickel-copper ratio is approximately 3 or 4 to 1.

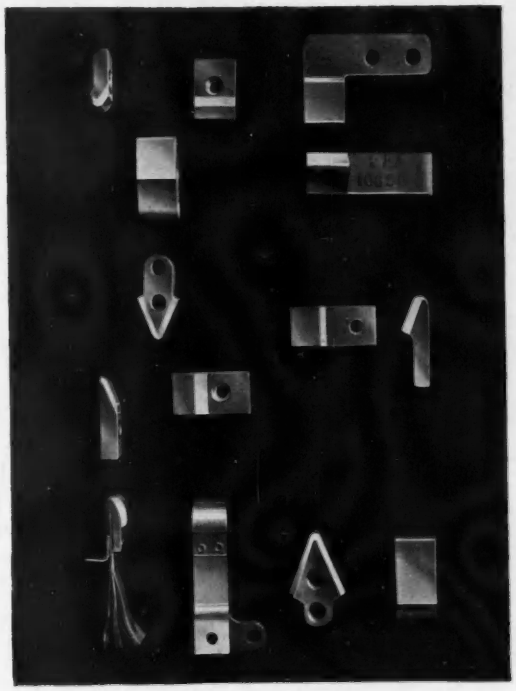


Fig. 22.—Arcing tips brazed on to copper or steel contacts.

Table VI and fig. 17 give the results of tests on a series of alloys containing 90 per cent tungsten but with various nickel-copper ratios.

The fact that the density does not fall away with the higher nickel-copper ratio alloys, although the strength does, discounts any suggestion that the decline in strength is due to porosity. It is evidently entirely governed by the composition of the copper-nickel-tungsten phase, and the higher the nickel-copper ratio, the more tungsten there will be in solution. Thus, from the mechanical strength point of view, it is an advantage to use a nickel-copper ratio of 3 or 4 to 1. Variations in the nickel-copper ratio with other tungsten contents show that, for each series, the highest strength is obtained when the ratio is 3 or 4 to 1, but the maximum strength is obtained when the tungsten content is about 90 per cent.

The effect of temperature on the properties of Heavy Alloy has been studied and methods have been developed for the production of the material in sheet or wire form.

INDUSTRIAL APPLICATIONS.

1. Radium Containers.

Heavy Alloy was originally developed to meet the need for a high density material in radium beam therapy.[14,15] The use of large masses of radium necessitates the provision of adequate protection for the clinical and nursing staff as well as for the patients. This protection is secured by heavy screening, but, at the same time, the actual container should be as small as possible if the radium

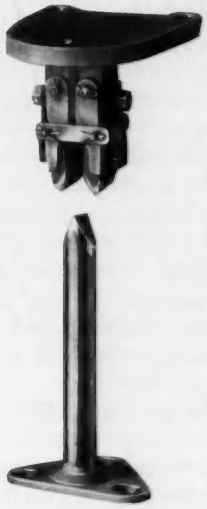


Fig. 23.—Arcing contacts of circuit breaker rated at 11/6.6 kV, 250 MVA.

is to be used efficiently, particularly in inaccessible positions. Roughly speaking, the absorption by metals of the harmful γ-radiation is directly proportional to their density and, for this reason, a small container can be used if it is made of a metal of high density. For many applications lead is probably ideal, as it can be cast easily and machined. Besides, it is very ductile and it is cheap. The density of lead, however, is only 11·4 gms./c.c. whereas that of Heavy Alloy is nearer 17 gms./c.c. and a very considerable reduction in the size of the container to give the same degree of protection is thus possible if Heavy Alloy is employed. The absorption coefficient, for hard γ-rays of a specimen of Heavy Alloy of density 16·5 gms./c.c., was estimated by inserting a slab 1 cm. thick in a narrow beam of radiation from radium filtered through 1 cm. of lead. The mean absorption coefficient over this range was found to be $\mu = 0{\cdot}83$ cm.$^{-1}$ With the same geometrical arrangement a slab of lead 1 cm. thick gave a mean coefficient $\mu = 0{\cdot}61$ cm.$^{-1}$ The ratio of the absorption coefficient of Heavy Alloy to that of lead is 1·37 which is close to the ratio of the densities, 1·32.

The distinct advantage in the use of Heavy Alloy instead of lead having been established, a number of applications was readily found for the material, and Heavy Alloy radium containers and nose pieces have now been adopted by most of the radium centres in Great Britain. Figs. 18 and 19 show a 5-gramme radium unit which has been constructed at the Radium Institute in London.[16] The radium is contained in twenty-five monel metal tubes, each containing 200 mgm. of radium element in the form of sulphate. The radium tubes are packed inside a steel bobbin which can be blown by air pressure from a lead storage-safe through a flexible tube into the Heavy Alloy nozzle.

Heavy Alloy radium containers are frequently of complicated construction, being fitted, for example, with interchangeable nose-pieces suitable for different applications.[17] In many cases, the design is such that the body of the container cannot be made from one piece of metal. A typical example of how a container can be built up by welding rings of Heavy Alloy is shown in fig. 20. A complete radium container although it may only be 6—8 ins. long and less than 6ins. in diameter will probably weigh as much as 100 lbs.

Other examples of Heavy Alloy as a screening material include its use for screens in X-ray tubes, for cosmic ray absorption measurements, for the construction of surface applicators in radium treatment and as a screening material in research work on radioactivity and radiographic examination of materials.

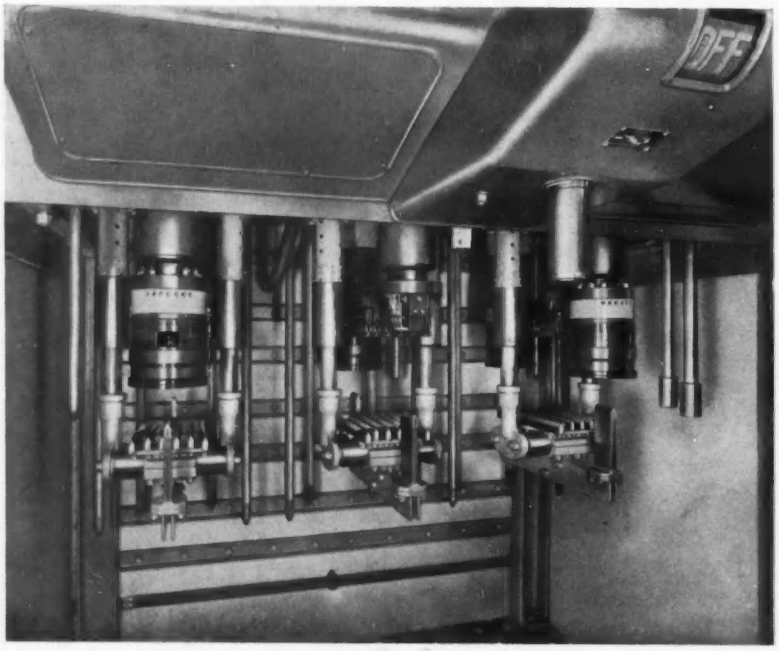


Fig. 24.—Arc control pots and main contacts fitted to a 1,000 MVA oil circuit breaker.

2. Electrical Contacts.

A great deal of Heavy Alloy is now used for the contacts of electrical circuit breakers. Its high melting point and consequent resistance to burning by the electric arc have made it a powerful aid in designing breakers of very high capacity.

In an alternating current oil circuit breaker an electric circuit is interrupted by the separation of one or more pairs of contacts under the surface of

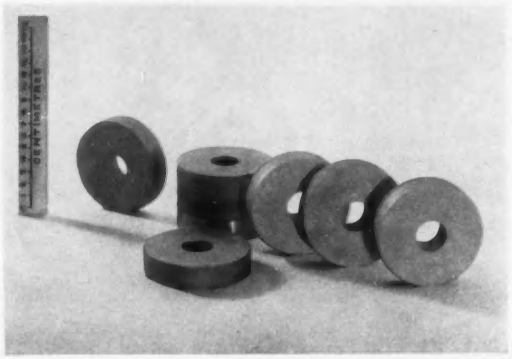


Fig. 25.—Heavy Alloy rings for use in a damper assembly for a marine reverse gear. The rings weigh approximately 3 lbs. each.

oil in a suitable container. The point or points of separation may, or may not, be surrounded by arc control pots or similar devices for controlling and aiding in the extinction of the ensuing arc. In either case the arc burns until it is extinguished by the blast resulting from the vaporisation of the oil by the heat of the arc.

The amount of energy liberated at the point where the arc touches the contacts is governed by the current and the cathode and anode voltage drops. For heavy currents, the sum of the two drops is about 45 volts, so that at each end of the arc about 22 kW is liberated for each 1000 A of current in the arc. As circuit breakers are commonly built for R.M.S. breaking currents of 10,000—50,000 amps., while breakers for 100,000 amps. have been constructed, the amount of heat produced is very large indeed. It goes, in part, to heating the contacts and, in part, to the evaporation and decomposition of the oil in the vicinity. The duration of the arc may be from 0·01 to 0·1 second according to the voltage of the circuit and the efficiency of the arc control device. Even this short time, however, suffices for the melting or vaporisation of some of the metal of the contacts.

According to the British Standard Specification for oil circuit breakers (B.S.S.116—1937) a circuit breaker must be capable of performing three "break" operations and two "make" operations in succession at its maximum short circuit capacity and still be in condition to make and carry its normal load current. It should be remembered that the peak making current may be 2½ times the R.M.S. breaking current. For this reason most breakers and especially those for very heavy breaking currents are provided with arcing contacts which make before, and break after, the main contacts which carry the normal load current. The burning is thus confined to the arcing contacts.

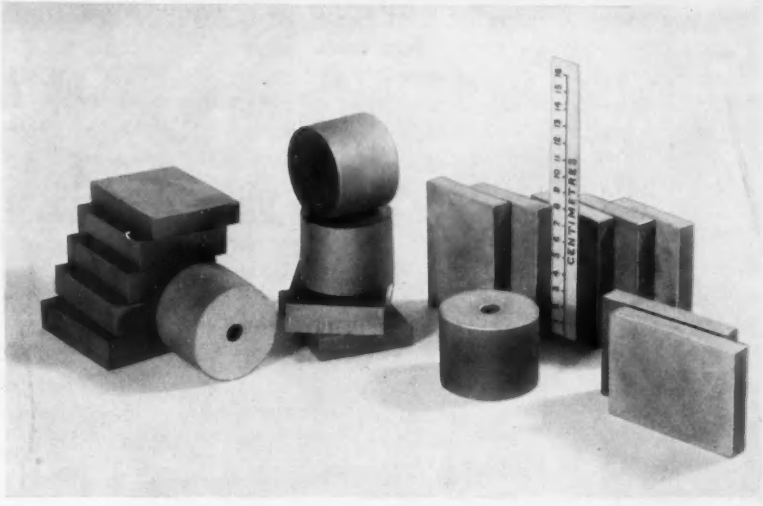


Fig. 26.—Heavy Alloy castings ready to be machined into vibration dampers.

Originally copper or brass was used for these arcing contacts. With the general increase in the severity of circuit breaker duty that took place some years ago, however, it was soon found that copper contacts were insufficiently resistant to burning and, in addition, were liable to weld together during the instant when the contacts were touching each other and carrying the short circuit current.

The more stringent requirements of service were reflected in the revision of B.S.S. 116 in 1937, by which the test duty cycle was altered from Break—Make Break to Break—Make Break—Make Break, resulting in an increase of approximately 100 per cent in the severity of the duty upon the contacts.

In the search for a more suitable material than copper, steel and one or two steel alloys were tried with moderate success before it was discovered that G.E.C. Heavy Alloy was ideal for the purpose. A comparatively thin layer of the alloy, $\frac{3}{32}$in. to $\frac{1}{4}$in. thick according to the circumstances, was found to be sufficient to limit the burning to a tolerable degree.

The alloy now in use for this purpose has the standard composition :—

tungsten	..	90 per cent
nickel	..	7·5 per cent
copper	..	2·5 per cent

The Heavy Alloy tips are pressed and sintered to the required shape and in the manner described above ; some of the types now being used are shown in fig. 21. The tips are attached to the contact stems by electrical brazing or silver soldering. The stem and tip are then finished to size by grinding. Figs. 22 and 23 show brazed contacts awaiting their final assembly. Special shapes of contact are sometimes built up by brazing together flat plates. A complete breaker in which both the fixed and moving contacts are protected by Heavy Alloy is shown in fig. 24.

Some appreciation of the reasons for the success of Heavy Alloy in this application may be gained by consideration of the following table :—

	Melting point	Boiling point
Tungsten	3400 degs. C.	5900 degs. C.
Nickel	1452 degs. C.	2325 degs. C.
Copper	1083 degs. C.	2310 degs. C.
Zinc	419 degs. C.	907 degs. C.

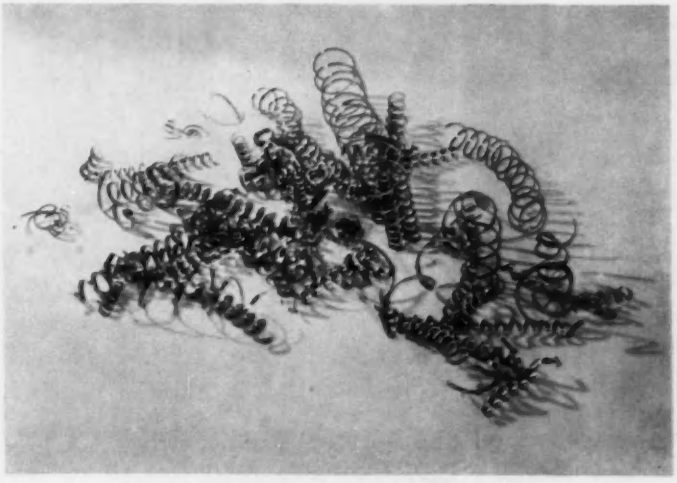


Fig. 27.—Turnings of Heavy Alloy.

The melting and boiling points of tungsten are over twice as high as those of copper, so that the erosion of the metal is correspondingly reduced. In addition, it appears that the admixture of copper and nickel has a beneficial effect as these metals vaporise at a lower temperature than that at which tungsten melts and therefore probably exercise a cooling effect upon the latter which, as can be seen by examination of a burnt contact, is left behind in a spongy condition. In this connection it is interesting to observe that brass gives better results than copper as an arcing contact material in spite of its low melting point of about 900 degs. C. Here again one constituent (zinc) vaporises at a temperature lower than that at which the other (copper) melts.

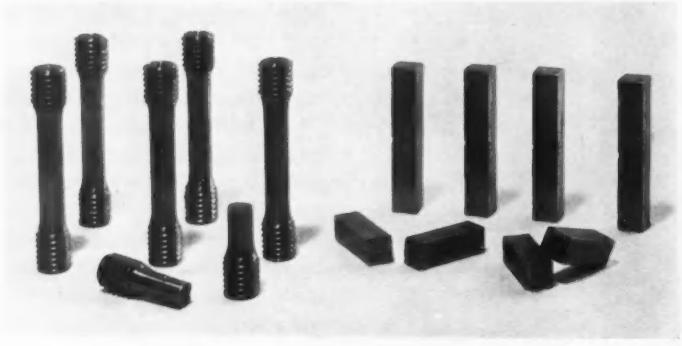


Fig. 28.—Machined test pieces.

3. Heavy Alloy as an Engineering Material.

The engineer is continually faced with the problem of balancing the moving parts of machines of all kinds. With the exception of the light alloys and lead, all engineering materials have very much the same density (Table I), and their use as balance weights is often inconvenient because of the volume required. Lead, which has a higher density, has a low tensile strength and a tendency to creep and thus is not a very suitable metal for many types of balance weights. Furthermore, its low melting point is a serious disadvantage to its use in a machine where there is any chance of overheating. The development of Heavy Alloy, with a tensile strength of 40 tons/sq. in. and a density over twice that of steel opens up new possibilities for the designer.

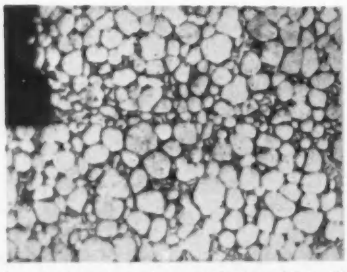


Fig. 29.—Photomicrograph of nickel weld in Heavy Alloy. The weld runs horizontally across from the angle of the black shadow on the left (magnification × 300).

It is an interesting fact that although the light alloys of aluminium and magnesium are playing such a leading part in modern aircraft construction, it is in this field that an ever increasing use is being made of Heavy Alloy.

For most engineering purposes, it is customary to supply the alloy in a form equivalent to castings from which the finished product may readily be machined. Figs. 25 and 26 show typical groups of such castings.

It is well known that tungsten is practically unmachinable, so that even if it were possible to obtain it in massive forms, it would be difficult to convert them into usable shapes. Heavy Alloy, although containing 90 or more per cent tungsten, is in marked contrast to this. The turnings shown in fig. 27, the test pieces in fig. 28 and the balance weights in fig. 31, illustrate its good machining properties.

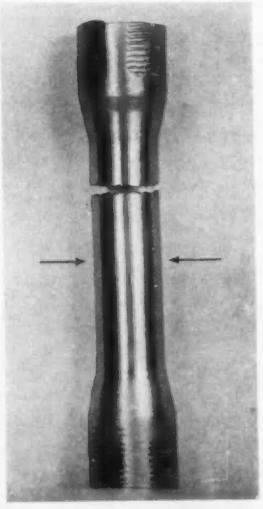


Fig. 30.—A test piece of Heavy Alloy welded at the section marked by arrows, failed as shown at 40 tons/sq. in.

Heavy Alloy may be brazed or silver soldered to other metals using the ordinary fluxes and a special method of welding in hydrogen has been developed for joining components of Heavy Alloy together, such as the radium container shown in fig. 19. This process takes advantage of the presence of the liquid phase when the temperature is sufficiently high and consists of heating up two or more pieces of Heavy Alloy with their planes of contact separated by a piece of nickel foil. If then the temperature of the whole mass is raised until the nickel melts, the molten nickel will alloy perfectly with the liquid phase of the Heavy Alloy and a homogeneous weld which is indistinguishable from the parent alloy will be

obtained, as is shown in the photomicrograph, fig. 29. As a further proof of the homogeneity of one of these nickel welds, a test piece was made up with a welded joint in the centre. The piece broke at nearly 40 tons/sq. in. at a point away from the weld as is shown in fig. 30.

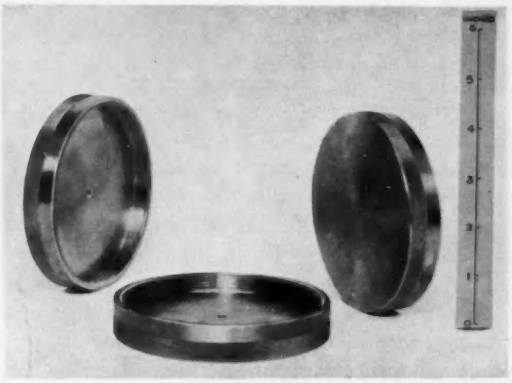


Fig. 31.—Machined balance weight caps.

Probably the major application for Heavy Alloy in engineering is as a material for mass balancing, but it is obviously not possible at the present time to discuss the nature of these applications. Since the moment of inertia of a rotating body is directly proportional to its mass, it follows that a piece of Heavy Alloy is capable of storing twice as much energy as a piece of steel of the same size and this again opens up further interesting applications.

There can be no doubt that when engineers and designers are accustomed to thinking of an alloy which has a density more than twice that of ordinary metals, the use of Heavy Alloy will be considerably extended.

REFERENCES.

1. C. J. Smithells, Nature 1937, *139*, 490.
2. Nickel Bulletin, 1937, *10* (5), 104.
3. G. H. S. Price, C. J. Smithells and S. V. Williams, J.Inst.Metals, 1938, *42* (1), 239.
4. R. Vogel, Z.anorg.allg.Chem. 1921, *116*, 231.
5. F. K. Gar'yanor and V. I. Tret'yakor, J.Tech. Phys. (U.S.S.R.), 1938, *8*, 1326.
6. F. H. Ellinger and W. P. Sykes, Trans. Amer. Soc. Metals, 1940, *28* (3), 619.
7. Freundlich, "Colloid and Capillary Chemistry," London 1926, p. 155.
8. W. Ostwald, Z. physikal. Chem. 1900, *34*, 495.
9. G. A. Hulett, Z.physikal. Chem. 1901, *37*, 385 and 1904, *47*, 357.
10. W. D. Jones, Metal Treatment, 1939, *5* (17), 13.
11. D. Hanson and W. T. Pell-Walpole, J.Inst. Metals, 1937, *61*, 265.
12. W. P. Sykes, J.Inst. Metals, 1938, *42* (1), 263.
13. W. D. Clark, J.Inst. Metals, 1940, *66*, 271.
14. J. C. McLennan and C. J. Smithells, J.Sci.Inst., 1935, *12* (5), 159.
15. L. G. Grimmett and J. Read, Brit.J.Radiology, 1935, *8*, 651.
16. L. G. Grimmett, Brit. J Radiology, 1937, *10*, 105
17. W. H. Rann, Brit. J. Radiology, 1938, *11*, 122.

Osram Fluorescent Tube Lighting in a Large Drawing Office.

Osram fluorescent tube lighting is particularly well adapted for the illumination of drawing offices, the low surface brilliance ensuring a virtually shadowless light which closely resembles north sky daylight. The tubes are designed for operation on A.C. mains of 200/250 volts and for almost instantaneous starting, while the efficiency is more than double that of a tungsten filament lamp.

Economic Factors of Power Production by Steam Plant.

By B. POCHOBRADSKY, M.I.Mech.E., M.Inst.F.,
Chief Engineer, Fraser & Chalmers Engineering Works.

THE measure of economy of power production is the total cost of a power unit (kWh) available for use outside the generating plant itself. In some cases there may be saleable by-products, for instance, ash; the net price realised for ash should be deducted from the cost of fuel. In other instances the steam turbine may, in addition to producing power, be supplying steam for some process work or heating. This combined output, power and heat, influences favourably the costs of both, and a correct apportioning of costs to each requires careful investigation. Our subject deals with power production only and we shall consider a simple plant unit consisting of boilers, steam turbine and electric generator with their auxiliaries.

The cost of power production might be conveniently divided into three groups:

(1) Materials.
(2) Wages and Salaries.
(3) Charges.

(1) *Materials.* Under this heading is understood everything that has to be supplied from outside sources to the generating station: fuel, lubricating oil, grease, replacement or repair materials, cleaning materials, water, chemicals for feed water treatment.

(2) *Wages and Salaries.* Payment to management, supervising staff, operating staff, maintenance staff, clerical staff, testing and research staff, labour.

(3) *Charges.* These would include interest on invested capital, depreciation, rent, rates, taxes, insurance, legal expenses, consultants' fees and any other expenses incurred but not belonging to group 1 or 2.

Let us examine the various items of costs assuming a given general price level.

(1) *Materials:* The most important item under this heading is the fuel.

Let N_{el} = output of the turbo-alternator in kW at its terminals;
$\overline{N}_{el}$ = Yearly average generated output of the turbo-alternator in kW;
$h/year$ = number of service hours per year.
$kWh/year$ = Total kWh generated per year.

From these definitions follows the equation:

$$\overline{N}_{el} = \frac{kWh/year}{h/year}$$

From tests (or guarantees) the hourly steam consumption M_1 in lbs./hour may be deduced for different outputs; it will be found that the hourly steam consumption plotted against the output is almost a straight line and therefore can be closely represented, during the time of service, by the equation:

$$M_1 = a + b\,N_{el}$$

This equation may be applied on the yearly average generated output and we obtain as yearly average of hourly steam consumption $\overline{M}_1$ in lbs./hour.

(1) $$\overline{M}_1 = a + b\,\overline{N}_{el}$$

where a and b are constants dependent on steam conditions and the efficiency of the turbine and alternator.

For the boilers, a similar relation exists between the hourly steam generation, M_1 lbs./hour, and the hourly fuel consumption, F lbs./hour.

$$F = c + d\,M_1$$

or applied on the yearly average of hourly steam consumption, we obtain the yearly average of hourly fuel consumption $\overline{F}$ lbs./hour:

(2) $$\overline{F} = c + d\,\overline{M}_1\,;$$

where c and d are constants dependent on steam pressure and temperature, feed water temperature, calorific value of fuel, efficiency of boiler plant and losses in the steam line.

Some power will have to be supplied by the turbo-alternator for driving all the auxiliaries of the whole plant:

Coal and ash handling plant, feed pumps, boiler fans, condenser pumps, alternator fans (if separately driven), light, ventilation, etc. The auxiliary power can be closely approximated by the expression:

$$m + n\,\overline{N}_{el}$$

m and n being constants for a given plant.

Let $\overline{N}_{eln}$ = yearly average net output (available for use outside of the generating plant unit).

then $\overline{N}_{eln} = \dfrac{kWh_n/year}{h/year}$ where kWH_n/year is the total net kWh per year available for use outside the generating plant unit.

It follows that: $\overline{N}_{el} = \overline{N}_{eln} + m + n\,\overline{N}_{el}$

or (3) $$\overline{N}_{el} = \frac{\overline{N}_{eln}}{1-n} + \frac{m}{1-n}$$

From equations (1), (2) and (3) we obtain :

$$\overline{F} = c + d[a + b\overline{N}_{el}] = c + d\left[a + b\left(\frac{\overline{N}_{eln}}{1-n} + \frac{m}{1-n}\right)\right]$$

$$= c + da + db\,\frac{m}{1-n} + \frac{db}{1-n}\,\overline{N}_{eln}$$

Let $A = c + da + db\,\frac{m}{1-n}$ and $B = \frac{db}{1-n}$

A and B are obviously constants for a given plant unit and they depend on steam pressure and temperature, temperature of condenser cooling water, efficiency of the plant unit, and calorific value of the fuel.

$$\overline{F} = A + B\,\overline{N}_{eln} \tag{4}$$

The fuel consumption in lbs. per net kWh_n (available for use outside of the generating plant unit) is :

$$f = \frac{A}{\overline{N}_{eln}} + B \tag{5}$$

Assuming that the cost of 1 ton (2,240 lbs.) of fuel delivered to the power station is P, the fuel cost per net kWh_n will be p :

$$p = f\,\frac{P}{2240} = \frac{P}{2240}\left(\frac{A}{\overline{N}_{eln}} + B\right)$$

The fuel cost per net kWh_n of a plant unit during the service time depends on certain constants characteristic of the plant, on the quality of fuel and on the yearly average net output.

Let us illustrate the equations (4) and (5) by an example of a 10,000 kW turbo-alternator with boilers and auxiliaries for which we assume the characteristic constants $A = 2600$ lbs./hour and $B = 1{\cdot}24$ lbs./kWh_n.

The yearly average of hourly fuel consumption during service time will be :

$$\overline{F} = 2600 + 1{\cdot}24\,\overline{N}_{eln} \tag{6}$$

the hourly fuel consumption at 10,000 kW net output will be $2600 + 12400 = 15{,}000$ lbs./hour and at no-load 2,600 lbs./hour.
The fuel consumption per kWh_n during service time at a net output $\overline{N}_{eln}$ will be :

$$f = \frac{2600}{\overline{N}_{eln}} + 1{\cdot}24 \tag{7}$$

Fig. 1 shows the equation (6) for $\overline{F}$ and (7) for f in our example. We see, in particular, the increase of fuel consumption per net kWh_n with decreasing net output ; the rate of that increase is comparatively small from the maximum net output down to about 40 per cent of that maximum ; for smaller outputs the rate of that increase grows rapidly. Differentiating the equation (7) we obtain :

$$df = -\,\frac{2600\,d\,\overline{N}_{eln}}{\overline{N}_{eln}^{\,2}}$$

and the rate of increase of fuel consumption for decreasing net output is :

$$\frac{df}{-\,d\overline{N}_{eln}} = \frac{2600}{\overline{N}_{eln}^{\,2}}, \tag{8}$$

which means that the rate of fuel consumption increase with decreasing net output is inversely proportional to the square of the net output.

We can also write :

$$\frac{-\,df}{d\overline{N}_{eln}} = \frac{2600}{\overline{N}_{eln}^{\,2}}$$

which means that the rate of decrease of fuel consumption with increasing net output is inversely proportional to the square of net output.

The fuel consumption per net kWh_n at $\overline{N}_{eln}$ is $\frac{2600}{N_{eln}} + 1{\cdot}24$, at 10,000 kW net output $\frac{2600}{10000} + 1{\cdot}24 = 1{\cdot}5$.

The increase in fuel consumption per net kWh_n at a net output $\overline{N}_{eln}$ over the fuel consumption at 10,000 kW expressed as a percentage of the consumption at 10,000 kW is :

$$\frac{\frac{2600}{\overline{N}_{eln}} + 1{\cdot}24 - \frac{2600}{10000} - 1{\cdot}24}{1{\cdot}5}\,100 = 173333\left(\frac{1}{\overline{N}_{eln}} - \frac{1}{10000}\right)$$

$$\frac{\triangle f}{f_{10000}}\,100 = \frac{173333}{\overline{N}_{eln}} - 17{\cdot}3333 \tag{9}$$

This equation is also plotted in fig. 1 and of course it also represents the increase of fuel cost per net kWh_n expressed as a percentage of the fuel cost per net kWh_n at 10,000 kW during service time.

Let us consider a power station consisting of five identical plant units, of which one is a standby only and the four units are in service during the year as follows :

Number of service hours per year	Average net output during service hours.
h_1	$(\overline{N}_{eln})_1$
h_2	$(\overline{N}_{eln})_2$
h_3	$(\overline{N}_{eln})_3$
h_4	$(\overline{N}_{eln})_4$

Using the equation (4) the yearly consumption of fuel F_y lbs./year for all machines during their service hours is :

$$F_y = [A + B\,(\overline{N}_{eln})_1]\,h_1 + [A + B\,(\overline{N}_{eln})_2]\,h_2 + [A + B\,(\overline{N}_{eln})_3]\,h_3 + [A + B\,(\overline{N}_{eln})_4]\,h_4 = A\,(h_1 + h_2 + h_3 + h_4) + B\,[\,(\overline{N}_{eln})_1\,h_1 + (\overline{N}_{eln})_2\,h_2 + (\overline{N}_{eln})_3\,h_3 + (\overline{N}_{eln})_4\,h_4]$$

Number of service machine hours per year :

$$h_m = h_1 + h_2 + h_3 + h_4$$

Total net kWh_n per year available for use outside the generating plant is :

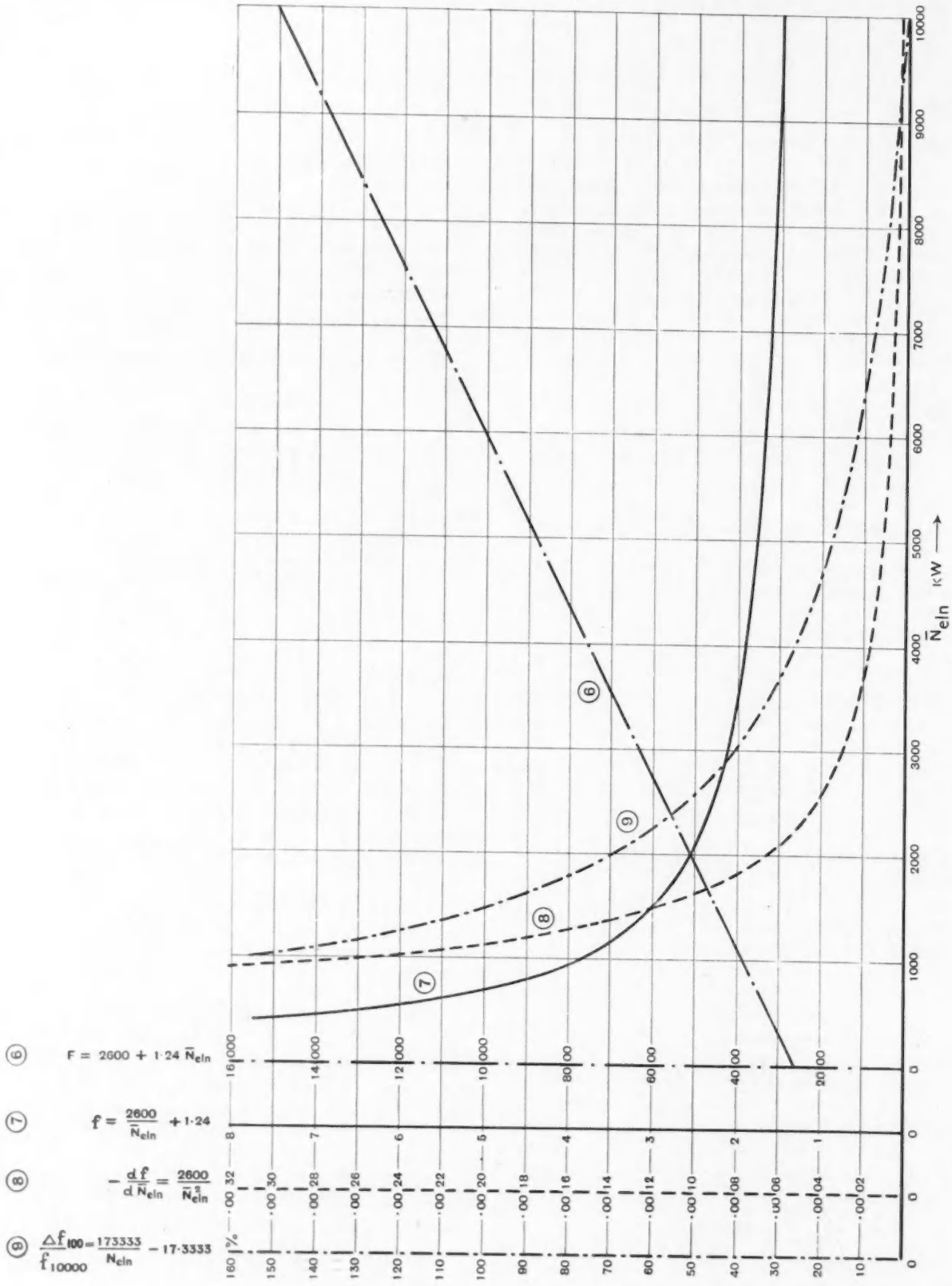

⑥ F = 2600 + 1·24 $\bar{N}_{eln}$
⑦ $f = \frac{2600}{\bar{N}_{eln}} + 1{\cdot}24$
⑧ $-\frac{df}{d\bar{N}_{eln}} = \frac{2600}{\bar{N}^2_{eln}}$
⑨ $\frac{\Delta f_{100}}{f_{10000}} = \frac{173333}{N_{eln}} - 17{\cdot}3333$
$\bar{N}_{eln}$ kW
%

Fig. 1.

$$kWh_n/year = (\overline{N}_{eln})_1\, h_1 + (\overline{N}_{eln})_2\, h_2 + (\overline{N}_{eln})_3\, h_3 + (\overline{N}_{eln})_4\, h_4$$

Therefore

$$(10) \qquad F_y = Ah_m + B\, kWh_n/year$$

As A and B are constants, we see that the total fuel consumption per year during service hours consists of two parts: one dependent on total machine hours per year, the other on the total net kWh_n/year. For a given total supply of kWh_n/year the fuel consumption during service hours will increase with increased machine hours; it follows that it is advantageous to reduce the machine hours; in other words, to reduce the number of plant units in service and to carry the load with the minimum number of plant units.

But if the service of a plant unit is not continuous but intermittent, then during the stoppages a certain amount of heat will be lost in the turbine, steam main and boilers; these losses may be reduced by good heat insulation and careful supervision. The duration of stoppages is likely to be regular; the rate of these heat losses during a stoppage is greatest at the beginning of the stoppage and decreases gradually. The total heat loss due to a stoppage grows with the duration of stoppage until the plant unit is cold.

As the rate of these heat losses is greatest at the beginning of a stoppage, it means that more frequent stoppages of shorter individual duration mean greater heat loss per year than less frequent stoppages of longer individual duration if the total duration of stoppages per year is the same in both cases. Further, frequent stopping and re-starting means greater fuel consumption for re-starting and paralleling operation. By methodical tests the best h_m will be established for a load variation over a day, week and year.

There will be blow-off losses due to stoppages and to load fluctuation; they grow with the number of stoppages per year and quick fluctuation of load and the degree of load fluctuation. Further, there will be blow-down losses which will depend on the quality of the feed water and the total quantity of feed water used per year; moreover, inferior quality of feed water may be the cause of a reduction of boiler and turbine efficiency and life. There might be loss of vacuum due to air leakages on the vacuum side of the plant or due to a dirty condenser; loss of vacuum means, of course, increased fuel consumption.

Assuming that the plant is kept in good working condition and that the service is regular, then there will be a certain supplementary fuel quantity per year F_{sy} lbs./year to cover the various incidental and unavoidable heat losses; or the total fuel consumption F_y' lbs./year is:

$$(11) \qquad F_y' = Ah_m + B\, kWh_n/year + F_{sy}$$

or the fuel consumption per net kWh_n averaged over a year is:

$$(12) \qquad f_y' = \frac{Ah_m}{kWh_n/year} + B + \frac{F_{sy}}{kWh_n/year}$$

Various factors are in current use: load factor, time factor, etc.; they are not always used in the same sense; let us therefore define:

$$\text{Plant load factor } l_p = \frac{kWh_n/year}{x\, N_{eln\ max}\, 8760}$$

Where x is the usual maximum number of plant units in service, $N_{eln\ max}$ the maximum net output of each plant unit (identical plant units are assumed).

$$\text{Service time factor } t_p = \frac{h_m}{x\, 8760}$$

Equation (12) becomes:

$$(13) \qquad f_y' = \frac{At_p}{l_p\, N_{eln\ max}} + B + \frac{F_{sy}}{l_p x N_{eln\ max}\, 8760}$$

In other words, the yearly average fuel consumption per net kWh_n consists of three parts: the first is proportional to the service time factor and inversely proportional to the plant load factor; the second is a constant; the third is inversely proportional to the plant load factor, but it should not be forgotten that F_{sy} depends to some extent on the service time factor; it is larger with a smaller service time factor. It can be said broadly that the yearly average fuel consumption per net kWh_n is not merely a function of the plant load factor, as is usually assumed, but a function of plant load factor and service time factor.

Equation (13) does not lend itself to easy analysis for practical purposes.

Let (14) $$\overline{N}_{eln}\, h_m = (\overline{N}_{eln})_1\, h_1 + (\overline{N}_{eln})_2\, h_2 + (\overline{N}_{eln})_3\, h_3 + (\overline{N}_{eln})_4\, h_4 = kWh_n/year$$

If one of the four plant units supplied a total of kWh_n/year in h_m hours its average output would be $\overline{N}_{eln}$.

Using equation (12) we obtain:

$$(15) \qquad f_y' = \frac{A}{\overline{N}_{eln}} + B + \frac{F_{sy}}{\overline{N}_{eln}\, h_m} = \frac{A}{\overline{N}_{eln}} + B + \frac{F_{sy}}{kWh_n/year}$$

and the fuel cost per net kWh_n p':

$$(16) \qquad p' = \frac{P}{2240} f_y' = \frac{P}{2240}\left(\frac{A}{\overline{N}_{eln}} + B + \frac{F_{sy}}{kWh_n/year}\right)$$

As F_{sy} contains banking losses which will be smaller with greater h_m, the third portion of fuel consumption per kWh_n decreases more quickly than in inverse proportion to h_m. The chief factor is $\overline{N}_{eln}$; the fuel consumption per kWh_n will diminish with increasing $\overline{N}_{eln}$ defined by equation (14).

But $\overline{N}_{eln}$ may be defined as follows:

(17) $\overline{N}_{eln} = l\, C_{max}$

l = load factor

C_{max} = net maximum capacity of a plant unit in kW. Equation (15) may be written:

(18) $$f_y' = \frac{A}{l\, C_{max}} + B + \frac{F_{sy}}{l\, C_{max}\, h_m}$$

The yearly average fuel consumption per net kWh_n will decrease with increased load factor (as defined by equation (17)) and with increased machine hours per year.

In the plant so far examined the plant units were identical; this is not the case as a rule. If the plant consists of plant units of different size and different efficiency then a simple fuel consumption formula is not obtainable and it may be advisable to try various combinations of plant units in a given load curve (daily, weekly, seasonal) to obtain the minimum total fuel consumption per year for a given total yearly supply of kWh_n/year. Supposing that, at most, four plant units are in service, then the total yearly fuel consumption in lbs./year will be:

(19) $F_y' = [A_1 + B_1 (\overline{N}_{eln})_1]\, h_1 + [A_2 + B_2 (\overline{N}_{eln})_2]\, h_2 + [A_3 + B_3 (\overline{N}_{eln})_3]\, h_3 + [A_4 + B_4 (\overline{N}_{eln})_4]\, h_4 + F_{sy}$

and the total yearly supply:

(20) $kWh_n/year = (\overline{N}_{eln})_1\, h_1 + (\overline{N}_{eln})_2\, h_2 + (\overline{N}_{eln})_3\, h_3 + (\overline{N}_{eln})_4\, h_4$

$A_1\ A_2\ A_3\ A_4$ and $B_1\ B_2\ B_3\ B_4$ are the characteristic constants for the four different plant units.

In making various combinations it might be found that the most efficient plant unit is not necessarily best suited for operation for the longest time in a year; for instance, if there is a small night load, it might be best covered by a smaller, though generally less efficient plant, because the smaller set might operate at a high load factor during the night, while a larger plant unit, though generally more efficient, would at an unavoidably low load factor show higher fuel consumption.

The equation (15), (16) or (18) applies generally to one plant unit or any number of identical plant units.

A plant unit is required to produce a certain maximum net output; the economy in fuel consumption improves with increased average net output (or increased load factor as defined by equation (17).) It is, therefore, desirable to organise the load so as to reduce the peaks; this is often possible in industrial concerns when certain machines might be stopped at the time when other unavoidable loads are put on without detriment to production. Reduction of peaks means also reduction in size, and therefore in the price of the plant itself.

The fuel consumption depends not only on certain characteristic constants and on average load, but on the maintenance of the plant in good working condition. In course of time efficiency, output and life of the plant might be reduced by poor quality of feed water causing corrosion, sediment on turbine blades or boiler tubes; falling steam pressure and temperature or reduced vacuum, due to dirty condenser or air leaks on the vacuum side of the plant, will cause loss in fuel and output. Defective insulation of the hot parts of the plant means increased fuel consumption (as an example, 1 per cent of heat lost in the steam main between the boiler and turbine means approximately 2 per cent increase in fuel consumption). Good maintenance will minimise all these losses; the operating and maintenance staff will be assisted in their endeavours by appropriate instruments and reliable continuous records of readings which "detect" the defects, or at least indicate that there is something amiss and so lead to a fruitful investigation. Efficient plant must be operated skilfully (not just operated) to give its inherent efficiency.

The other items under the heading of "Materials" represent comparatively a small amount, and their yearly costs increase generally with the total net kWh_n/year. They may be expressed by:

$$D + E\, kWh_n/year$$

where D and E are constants.

Lubricating oil deserves special attention; viscosity of oil higher than necessary, whether due to "heavier" brand of oil or to too low bearing temperatures, causes only higher bearing losses; dirty oil increases wear of various parts of the plant; oil leakage is not only a waste, but may constitute a fire hazard.

The cost of materials for repairs or replacement will naturally be reduced by careful supervision.

(2) *Wages and Salaries.*

For a given plant and mode of operation these may be constant or slightly increasing with the total net kWh_n/year; they may be conveniently represented by:

$$G + H\, kWh_n/year$$

where G and H are constants, H being in most cases small or even zero.

(3) *Charges.*

Provided that there is no change in the rate, the charges will represent a constant amount per year independent of the total yearly supply of energy.

The various items of the charges will, as a rule, depend on the cost of the plant, the insurance being the only item which may cover not only the plant, but also the personnel. They will be governed by individual circumstances; interest and depreciation deserve special consideration.

The capital cost of the plant will be provided either by loan or by investment of new capital; in

either case the plant is expected to earn the interest (or dividend) on the capital cost.

The yearly amount of interest (or dividend) on the capital cost is :

$$(21) \qquad I = C\,\frac{p_1}{100} = C(r_1 - 1)$$

where C = capital cost

p_1 = rate of interest in per cent

$$r_1 = 1 + \frac{p_1}{100}$$

As the yearly amount for interest is proportional to the rate of interest, it is advantageous to use for capital cost the cheapest possible capital. Should there be an opportunity to invest available capital so that it would earn a higher rate of "dividend" than the rate of interest demanded on a loan, then such loan should be used for the capital cost.

The yearly amount of depreciation should be such that at the end of the "economic" life of the plant the original cost is paid off. The various parts of a power station may have different "economic" life. In Great Britain, 60 years for land, 30 years for buildings, and 20 years for plant are usually assumed.

Let $C = C_L + C_B + C_P$

where C_L = cost of land

C_B = cost of buildings

C_P = cost of plant

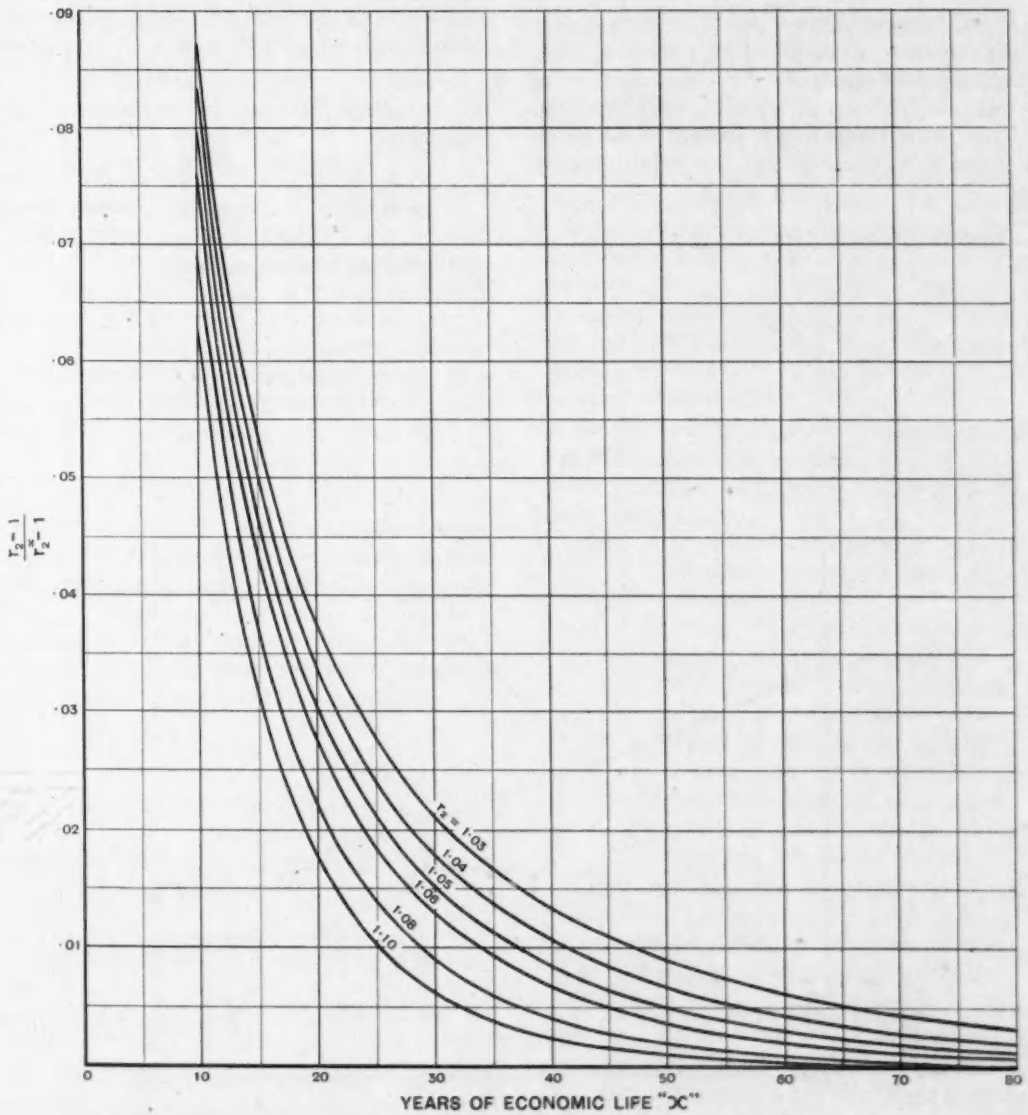


Fig. 2

Let x_L = economic life in years of land
x_B = economic life in years of buildings
x_P = economic life in years of plant
p_2 = rate of interest in per cent for depreciation

$$r_2 = 1 + \frac{p_2}{100}$$

The yearly amount for depreciation for the power station is :

$$(22) \qquad D_S = C_L \frac{r_2 - 1}{r_2^{x_L} - 1} + C_B \frac{r_2 - 1}{r_2^{x_B} - 1} + C_P \frac{r_2 - 1}{r_2^{x_P} - 1}$$

In fig. 2 the expression $\frac{r_2 - 1}{r_2^{x} - 1}$ is shown for different rates of interest $p_2 = (r_2 - 1)\ 100$ and for different numbers of years of "economic" life.

From the formula (22) and fig. 2, it is clear that the yearly amount of depreciation decreases with increasing years of economic life ; further, it is lower with a higher rate of interest ; therefore, the yearly depreciation amount to be earned by the plant will be lower if it is invested at the higher rate of interest.

Fig. 3 shows the factor $r - 1 = \frac{p}{100}$ in equation (21) and $\frac{r-1}{r^{20} - 1}$ a factor in equation (22), for 20 years of economic life for varying rates of interest.

For instance, should $p_1 = 5$ per cent
$I_P = C_P \times \cdot 05$
and should $p_2 = 8$ per cent

$D_P = C_P \times \cdot 02185$ for 20 years of economic life and $I_P + D_P = C_P \times \cdot 07185$ or 7·185 per cent of the capital cost C_P.

If we consider the plant generally, the yearly amount of interest and depreciation is :

$$I_P + D_P = C_P (r_1 - 1) + C_P \frac{r_2 - 1}{r_2^{x_P} - 1}$$

Should $r_1 = r_2 = r$ (or $p_1 = p_2 = p$)

$$\text{then } I_P + D_P = C_P(r-1)\left(1 + \frac{1}{r^{x_P} - 1}\right) = C_P(r-1)\frac{r^{x_P}}{r^{x_P} - 1}$$

Equation (21) can be written for the power station :

$$I_S = C_L (r_1 - 1) + C_B (r_1 - 1) + C_P (r_1 - 1)$$

Using this equation and equation (22), the interest (or dividend) and depreciation will require the following amount per year :

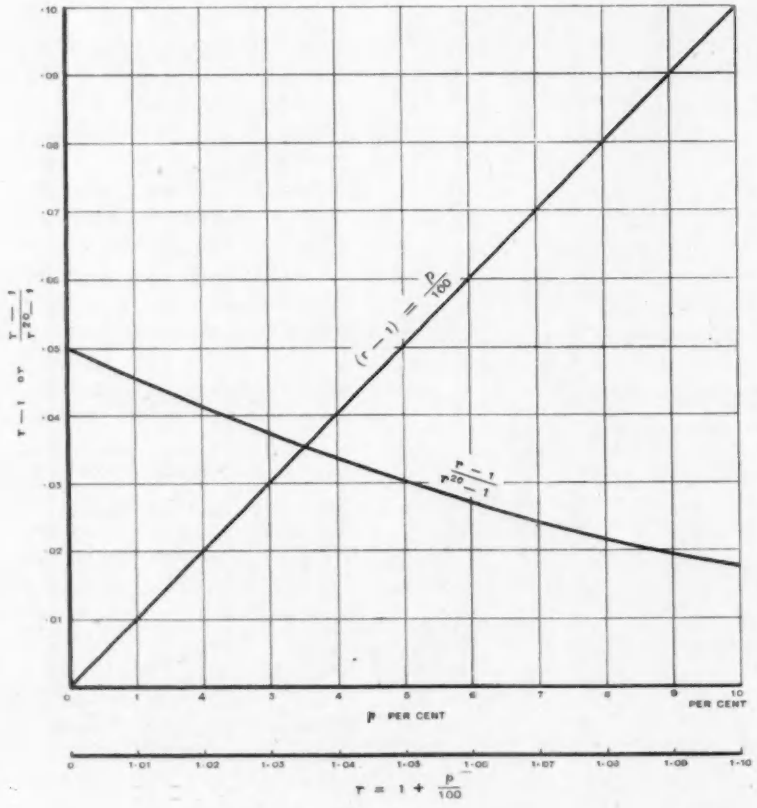


Fig. 3

$$I_S + D_S = C_L\left[r_1 - 1 + \frac{r_2 - 1}{r_2^{x_L} - 1}\right] + C_B\left[r_1 - 1 + \frac{r_2 - 1}{r_2^{x_B} - 1}\right] + C_P\left[r_1 - 1 + \frac{r_2 - 1}{r_2^{x_P} - 1}\right]$$

Let us examine the factor in brackets for the plant portion:

Assume $p_1 = 4$ per cent, $p_2 = 6$ per cent, $X_P = 20$ years, then $r_1 = 1{\cdot}04$, $r_2 = 1{\cdot}06$

and the factor is: $1{\cdot}04 - 1 + \dfrac{1{\cdot}06 - 1}{1{\cdot}06^{20} - 1} = {\cdot}04 + {\cdot}02718 = {\cdot}06718 = \dfrac{6{\cdot}718}{100}$

In other words, the factor represents 6·718 per cent of C_P.

The three factors in brackets may be designated as: p_L per cent, p_B per cent and p_P per cent and we obtain

$$I_S + D_S = C_L \frac{p_L}{100} + C_B \frac{p_B}{100} + C_P \frac{p_P}{100}$$

and this calculated total may be put $= C \dfrac{p_S}{100}$ where p_S is the percentage capital charge on the total cost of the power station.

Any other items under "Charges" are incurred either once (at the time of erection of the power station), in which case they can be treated in exactly the same way as the cost of the power station, or they are incurred yearly and can simply be added to the yearly amount of interest and depreciation. It is clear that the total yearly charges Ch can be expressed as a percentage p_C (in per cent) of the capital cost:

$$(23) \qquad Ch = C \frac{p_C}{100}$$

It has been seen that the fuel consumption (or cost) in a power station consisting of a number of plant units of different size and efficiency does not lend itself to ready analysis; hence to obtain a simple guide, let us assume one plant unit, or a power station consisting of a number of identical plant units to which equation (16) is applicable:

$$p' = \frac{P}{2240}\left(\frac{A}{\overline{N}_{eln}} + B + \frac{F_{sy}}{kWh_n/year}\right)$$

The yearly cost of all other materials $= D + E\,kWh_n/year$

of wages and salaries $= G + H\,kWh_n/year$

of charges $= C \dfrac{p_c}{100}$

or the yearly average cost of a net kWh_n of the last three items is:

$$\frac{D + G + C\dfrac{p_c}{100}}{kWh_n/year} + E + H$$

or the total yearly average cost of a net kWh_n including fuel cost is p_t:

$$(24) \quad p_t = \frac{P}{2240}\,\frac{A}{\overline{N}_{eln}} + \frac{1}{kWh_n/year}\left[\frac{F_{sy}P}{2240} + D + G + C\frac{p_c}{100}\right] + E + H + \frac{P}{2240}B$$

or the yearly average production cost of a net kWh_n consists of three parts:

1st part $\dfrac{P}{2240}\,\dfrac{A}{\overline{N}_{eln}}$ which is inversely proportional to the yearly average net output $\overline{N}_{eln}$ of a plant unit.

2nd part $\dfrac{1}{kWh_n/year}\left[\dfrac{F_{sy}P}{2240} + D + G + C\,\dfrac{p_c}{100}\right]$ which is inversely proportional to the total yearly supply of energy kWh_n/year.

3rd part $E + H + \dfrac{P}{2240}B$ which is independent of the load and of the total supply of electrical energy.

p_t is, of course, expressed in the same money units at P, the price of a ton of fuel.

Should the cost of a net kWh_n after transformers, or after transformers and at the end of a transmission line, be required, then of course the fuel cost has to be increased in inverse ratio to the yearly average transformer (and transmission line) efficiency; the other materials, wages and salaries and charges for transformers and transmission line, buildings and land, would be treated in a similar manner as for the power station itself.

It has been shown that the production cost of a net kWh_n is affected by many factors; equation (24) gives a number of them and these may be divided into several classes:

(*a*) A, B, F_{sy} which represent cycle and plant efficiency, quality of fuel and operation efficiency.

(*b*) $\overline{N}_{eln}$ and $kWh_n/year$ which represent load and operation organisation.

(*c*) D and E which indicate price and quality of some of the materials.

(*d*) G and H indicating the rates of wages and salaries, and the standard of the staff.

(*e*) C cost of the plant or station.

(*f*) P price of fuel, p_c rate of interest, rates, taxes and other local factors.

But to obtain the equation (24) a number of assumptions has been made; some of them are close approximations, but one stipulates that the plant is to be maintained in good working order.

It will be clear that the study of all these factors affecting the economy of power production can be most fruitful by collaboration between the purchaser and the manufacturer of the plant and indeed, in many instances, between the purchaser of the plant and his customers, the buyers of electrical energy.

Brightness and Contrast in Illuminating Engineering.

By R. G. HOPKINSON, Ph.D., A.M.I.E.E., F.R.P.S., W. R. STEVENS, B.Sc., F.I.E.S., and J. M. WALDRAM, B.Sc., F.Inst.P., F.I.E.S.
G.E.C. Research Laboratories.

This paper was read before the Illuminating Engineering Society on 10th December, 1940, and is reprinted from the Journal of that Society.

The nature of contrast, and the part which it plays in the recognition of objects is discussed. Contrast is subjective, and cannot be directly measured, but there apparently exists some general relationship between the photometric brightnesses of objects, the general brightness of the scene and the corresponding impressions received.

Some curious visual effects which can be seen in the blackout are explained, and the part played by contrast in some everyday problems of seeing is illustrated by examples.

1. CONTRAST AND RECOGNITION.

A DISCUSSION of "contrast" necessarily involves the whole process of seeing, for whatever may be its precise meaning, "contrast" is recognised as one of the factors which help to make things visible, and its interest lies in its ability to do so.

We tend to think of the eye as merely an apparatus for the perception of light; and much of the discussion of the eye in technical literature has led the technical reader to this conception, partly as a result of the necessary restriction of experiments to simple cases. But the process of seeing is infinitely more subtle than the perception of light. It was once remarked that "the eye is a very bad photometer"; but it would have been wiser to have said that "the photometer is a poor gauge of seeing," for seeing is the gathering of intelligible information from the light reaching the eyes, and it stands in much the same relationship to perceiving light as does the appreciation of music to the perception of sound. We have studied the ability of the eyes to perceive fine detail, low brightnesses, small colour differences, and so forth; but the step from these measurements to the use of our eyes in everyday seeing is very much like the step from measurements of the perception of loudness and timbre to the understanding of music or poetry.

Seeing does not even depend upon the perception of light alone. For example, in a thick fog by day there is plenty of *light* to be perceived, but we remark "I could not *see* a *thing*!" It is "things" not light, which we are interested to see; and we cannot see in a fog not because there is no light, but because there are no darknesses to make a recognisable pattern. But a kaleidoscopic pattern of light and dark and colour, which presents plenty of material for the eye to perceive, may give no *information* at all (fig. 1); and a few lines (fig. 2) may convey quite a definite impression of a face, with a good idea of expression, age, sex, character, etc., even though the lines have made no attempt correctly to represent contrasts, colour, or many of the important details of the face. It seems that we see not by light alone, nor by darkness alone, nor by contrast, nor colour alone, but by recognising the familiar geometry of the brightness and colour contrasts presented to us. Even a crude indication of this geometry may be sufficient to give a great deal of information either directly or by inference. Recognition thus depends upon:

(1) the amount of contrast, and colours;
(2) the geometry of their boundaries and distribution;
(3) familiarity with the scene.

Fig. 1.—Contrasts conveying no information.

2. MEANING OF "CONTRAST."

The term "contrast" is used in many senses other than the visual, and its use gives the key to its meaning. We speak of the contrast in music, between say a quiet and delicate passage for woodwind and a succeeding fortissimo passage with the full orchestra: of the contrast of character between say two members of the same family: of the

comical contrast of a short fat man beside a tall thin man, and so on. In all these cases, as well as in contrasts in brightness and in colour, contrast means "*the sensation of difference between juxtaposed things.*" The contrasting things may be juxtaposed in time or in space, as the above examples show; visual contrasts generally occur between objects juxtaposed in the field of view. The important point in the definition is that contrast is a *sensation* of difference. The sensation must arise from a physical difference in photometric brightness or in spectral distribution, but it is not identified with the physical difference; the same physical differences may produce in various circumstances "sensations of difference" (*i.e.*, contrasts), which are not the same. On the other hand some kinds of physical difference (for example of spectral distribution) may produce no contrast at all; this is what occurs when a trichromatic colorimeter is balanced.

Since contrast is a sensation, it is resident in the observer and it cannot be measured directly. All that can be directly measured are the brightnesses and spectral distributions which bring it about; the contrast itself must be gauged from these measurements by experience of similar situations or otherwise (see Section 6). There is, however, one quantity sometimes known as contrast in which the word has a specialised meaning, which must not be confused with the sense in which it has been used above. The photographer uses "contrast" to express a characteristic of sensitive materials and processes, and in this sense it is a definable and measureable physical quantity. Broadly speaking, it represents the relationship between a brightness ratio in the original scene to the corresponding brightness ratio in the reproduction of that scene. It is a measure of the degree to which the material or process over- or under-emphasises brightness differences in the original. It is generally expressed as the slope of the curve of density plotted as a function of the logarithm of the exposure, and is often known as "Gamma."

3. RANGE OF BRIGHTNESS RESPONSE.

The enormous range of brightnesses to which the eye will respond is often pointed out: in problems of vision in the blackout it is necessary to consider brightnesses below one millionth of an equivalent foot-candle; in a sunlit snow scene the brightness may be over 10,000 equivalent foot candles. The absolute overall range is about 10^{12} : 1. But the eye cannot cover the whole of this enormous range *at once*, and it is much more important in practical seeing to consider what may be called the instantaneous range of the eye. There is an analogy with an electrical test set, which may be a milliammeter provided with a set of shunts. The unshunted instrument may be able to measure satisfactorily currents as low as 0·1 milliampere; with the largest shunt in the range it may be able to measure 100 amperes, so that its total useful range is 1 million to 1. But with any one shunt its useful range is only about 20 : 1, with the 10-amp. shunt connected, variations of 10 milliamps will not be indicated, while currents of 100 amps. will damage the instrument. The eye functions very similarly; it is like a milliammeter or oscilloscope which automatically selects a shunt which will bring

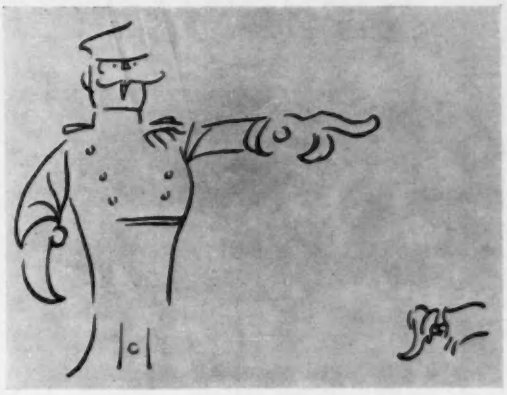


Fig. 2.*—Information conveyed by geometry alone.

conveniently on the scale the whole of a range of varying currents applied to it. It cannot, however, deal with too great a variation at once, and if an excessive variation is applied to it, then the maxima will be off the scale and will cause damage, or the minima will not be distinguishable from zero. In much the same way a radio receiver with automatic volume control adjusts its gain automatically to suit the level of the incoming signals; but exceedingly strong or very weak signals may lie outside the capabilities of the a.v.c. circuit and the output will be distorted.

These two analogies illustrate two phenomena of the eye which are important in seeing. The first is that the instantaneous range of the eye is limited, and it cannot simultaneously respond to very high and to very low brightnesses, or deal with one brightness of small extent which is very different from the general level to which the eye is adapted. The second is that as the general level of brightness of the scene is reduced, there comes a point when the darker parts of the scene fail to register as sensations; the eye has adapted as far as it can, but they are still too dark to be seen. Moreover, as the general brightness of the scene is reduced, relative brightness differences which are clearly distinguishable at the higher brightness levels become much less visible; the sensitivity of the eye to small differences of brightness is reduced so that a given relative change of brightness produces less sensation.

*This sketch by FOUGASSE is reproduced from "*Fun Fare*" by permission of Messrs. Hutchinson & Co. (Publishers), Ltd.

For example, when the eye is adapted to a brightness of 1000 e.f.c., brightnesses below 5 e.f.c. appear black, although when dark-adapted a brightness of 5 e.f.c. appears uncomfortably bright. When one

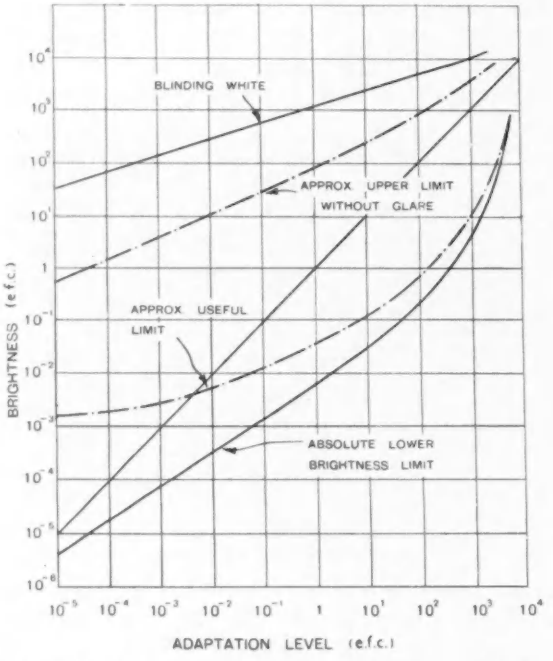


Fig. 3.—Instantaneous visual range related to adaptation level.

is out of doors on a bright sunny day, the eye is adapted to a brightness of about 1000 e.f.c., and little can be seen through an open door of the objects in a room indoors. The prevailing brightness inside the room may be of the order of 10 e.f.c., so that all but the brightest objects will be darker than the lower limit of brightness at the level of 1000 e.f.c. If one enters the room, the eye adapts to the average level of 10 e.f.c., and the objects in the room become distinguishable, for their brightnesses are within the useful range to which the eye will now respond. Most of the objects outside the window can also be distinguished, because they are still within the range of the eye adapted to 10 e.f.c., but it may be found that white clouds, seen through a small window, appear as glaring white masses devoid of detail, because the brightnesses here are beyond the upper limit for an average brightness of 10 e.f.c.

Similarly, when one leaves a brightly lighted room for the conditions of the blackout, hardly any detail can be resolved until the eye has had time to adapt to the very much lower brightness. In this case, however, the eye has to work under conditions near to its absolute lower limit; it can see very little of brightnesses less than the adapting brightness. Many of the objects in the field of view are below this lower limit, and hence no matter how long the eye is given to adapt, these objects will remain indistinguishable.

The curves of fig. 3 illustrate some of these phenomena. They represent the greatest and the least brightnesses to which the eye will simultaneously respond, when it is adapted to different brightnesses which are plotted as abscissae, and which are also represented by the straight line drawn at 45°. The intercept between the two curves shows the instantaneous range of the eye, and the relative positions of the ordinates on the two curves and the ordinate on the straight line shows the manner in which the range of the eye lies about the adaptation level. It will be seen that at the higher levels the range is great and is more or less symmetrically disposed about the adaptation level; *i.e.*, the eye can conveniently distinguish brightnesses which are both brighter and darker than the general adaptation level. But when the adaptation level is reduced the range becomes less, and it is no longer symmetrically disposed about the adaptation level; the eye can distinguish brightnesses greater than that to which it is adapted, but not much of brightnesses less than this.

4. EFFECTS OF RANGE IN PHOTOGRAPHIC REPRODUCTION.

The adaptation characteristics of the eye determine the laws of photographic reproduction. In order to represent a scene faithfully, it is necessary for the brightnesses of the reproduction to give the same sensations of brightness under the given viewing conditions as did the original brightnesses which they represent. The exact mechanism, as far as it is known at present, has been described in other communications, and need only be recapitulated here.

Whilst the total range of the eye is of the order of 10^{12} to 1, the range of a good photographic paper is of the order of only 50–1, and of a good transparency of the order of 10^3 to 1. It is therefore quite impossible to represent on photographic paper the total absolute range of the eye.

This is, however, unnecessary. At the most, the photographic paper need cover only the "instantaneous" range of the eye. Clearly if an observer cannot distinguish detail inside a room when he is out-of-doors, there is no need for a photograph to do so.

The "instantaneous" range of the eye is of the order of 1000 to 1, which is still greater than that of a photographic paper: but the range of brightness of the important parts of most *scenes* is seldom greater than about 30–1, so that these limitations of the photographic paper are not always such a drawback as would appear.

Provided that the range of brightness of the scene can be represented on the photographic paper or transparency, the proper densities of the photo-

graph are then determined by the adaptation levels of the eye both when viewing the scene and when viewing the print. The practical effect of these laws of photographic reproduction are well known. A given print viewed under changing viewing conditions may give a changing impression; for example, a photograph of a sunlit scene when viewed in bright artificial light will, if the print densities are correctly chosen, give a faithful impression of the original brilliance of the scene. If the brightness of the viewing conditions is much reduced, there comes a point where the sunlight has an unreal, watery appearance, and as the brightness is further reduced, the scene appears very like a scene by moonlight. Most photographers are well aware of this phenomenon, and have, for example, to make a mental allowance for the difference in appearance of a print in the darkroom and its appearance in the light. Fortunately the range of illuminations, over which the print gives a faithful impression, is large—roughly from 5 to 500 ft. candles in ordinary surroundings. When the illumination falls below 5ft. candles, the correctness of a given representation will depend closely upon the illumination.

5. SUBJECTIVE OR APPARENT BRIGHTNESS.

The problems of photographic reproduction demonstrate in the most arresting manner that the brightness of an object as measured by a photometer often bears little relation to the impression of this brightness on the observer. A brightness measured

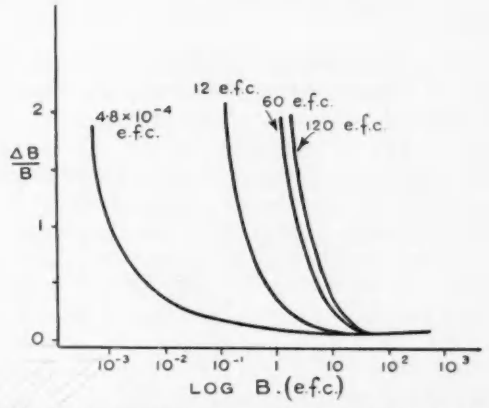


Fig. 4.—Brightness discrimination sensitivity for three adaptation levels (Abribat).

with a photometer must therefore be interpreted in terms of the general adapting conditions under which the brightness will be seen. This is usually done automatically as a result of experience, but confusion arises in unfamiliar problems or when an attempt is made to interpret brightness values in strict relationship to their photometric value.

Some system of interpretation of brightness values is therefore essential. The general consensus of authoritative opinion is, however, that sensations are not scalar or measurable, and consequently no strict relationship between physical brightness stimulus and sensation of brightness can be obtained. Some general relationship does, however, exist, and attempts to measure it have met with some success. For example, over a limited range equal increments of sensation appear to be produced by equal ratios of brightness stimulus. This is the well known Weber-

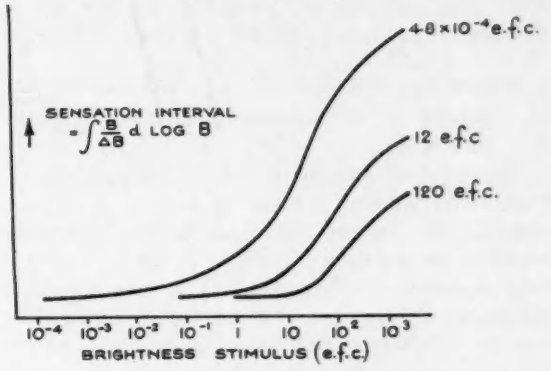


Fig. 5.—Sensation-stimulus relationship (Abribat).

Fechner Law. This law by no means expresses the complete relationship between brightness stimulus and sensation, however, and as a result attempts have been made to explore this complete relationship. Of these attempts, those of Abribat and of Pitt yield data which are in general agreement, although the methods employed were different.* Abribat obtained a series of curves relating brightness discrimination sensitivity with the brightness of the test field, for different values of adaptation level (see fig. 4). From the curves he derived, by integration, the family of curves shown in fig. 5, in which the brightness of the test field is correlated with the "sensation" of brightness produced. The validity of this integration, which was first performed by Fechner, has been questioned, but Abribat's ordinate scale of "sensation" can be reasonably replaced by an arbitrary scale of "apparent brightness." Pitt used a method of binocular matching in order to obtain his data; in this case a direct relationship between the apparent brightness and the physical brightness was obtained, without the need for any intermediate steps. Hopkinson has obtained a set of data using a three-part field, in which two parts formed the two comparison brightnesses, and the third part the adapting field; this method involved the assessment of large steps of brightness difference instead of just perceptible steps. The data obtained showed the same general tendencies as those of Abribat and of Pitt.

*M. Abribat. Réunion de l'Institut d'Optique, 1935.
F. H. G. Pitt. *Proc. Phys. Soc.* Vol. 51, p. 810, 1939.

A family of curves, which has been derived from a study of all the available data, appears to give a useful approximation to the characteristics of the eye (fig. 6). These curves show that at any given adaptation level the Weber-Fechner law holds over a fairly limited range (*i.e.*, those portions of the curves which are straight lines) and that at both high and low brightness levels equal ratios of stimulus no longer produce equal differences of sensation. This is, of course, in accord with practical experience. With such a family of curves, the problem of the quantitative study of contrast can be attacked.

6. QUANTITATIVE STUDY OF CONTRAST.

(*a*) *Contrast as a function of apparent brightness difference.*

In order to employ the data of fig. 6 for the quantitative study of contrast, it is assumed that the sensation of contrast produced by two contrasting brightnesses is a direct function of the difference in their apparent brightnesses. This assumption is reasonable for contiguous areas, but it may not be true for widely separated areas. Most problems are, however, concerned with contiguous areas.

Care has to be taken in using the numerical scale of apparent brightness. The scale is quite arbitrary, and has been so chosen that one unit represents a just perceptible brightness difference or contrast. It cannot, however, be assumed from this that a contrast of four units appears four times as great, or that in general, the scale is strictly additive. The numerical value of a contrast is an indication of the general order of that contrast. Neither will the scale necessarily discriminate accurately between two contrasts having nearly the same numerical value.

(*b*) *Method of assessing contrast.*

Most problems of contrast in illuminating engineering enquire either (*a*) whether the given contrast is greater or less than another contrast, or (*b*) whether the given contrast is adequate for the visual task. Both these forms of the problem can, in many cases, be attacked using the data of fig. 6. Three values of brightness need to be known, *i.e.*, the values B_1 and B_2 of the contrasting brightnesses, and the brightness B_3 to which the eye is adapted, which determines the curve in fig. 6 to be used. Of these, B_1 and B_2 can be measured with a photometer, but the determination of B_3 is more difficult. Fortunately, as will be seen, for the order of accuracy of which the method is capable, the determination of B_3 to within ± 100 per cent is usually sufficiently accurate.

In order to determine B_3, either an area can be chosen by inspection which appears to be of the average brightness of the scene, and the brightness of this area measured with a photometer; or else the vertical illumination falling on the eye of the observer (any very bright sources of light being screened off) can be measured, and the average brightness of the scene calculated from the value obtained.

For example, suppose that the brightness of an object in a lighted street under peace-time conditions is 0·05 e.f.c., and the background behind it is 0·5 e.f.c. The average brightness of the scene is of the order of 0·1 e.f.c. From the curve corresponding to 0·1 e.f.c. it is seen that at this adaptation 0·05 e.f.c. corresponds to an apparent brightness of 6 on the arbitrary apparent brightness scale, whilst 0·5 e.f.c. corresponds to an apparent brightness of 23. The contrast between these two objects is then taken as $23-6=17$ arbitrary units. Now consider the same conditions reduced by 1 : 1000, which approaches blackout conditions. B_3 is then 0·0001 e.f.c., B_1 is 0·00005 e.f.c., and B_2 is 0·005 e.f.c. From the curve corresponding to 0·0001 e.f.c. it is seen that at the new adaptation B_1 (0·00005 e.f.c.) has an apparent brightness of 0, and B_2 (0·0005 e.f.c.) has an apparent brightness of 2, the contrast now being $2-0=2$ arbitrary units. This is very much less than the contrast under the brighter conditions, a result which agrees with experience.

In order to show that the estimation of B_3 is not critical, suppose that in the first case a 100 per cent error had been made, and B_3 had been estimated to be either 0·2 e.f.c. or 0·05 e.f.c. In the first case the contrast between B_1 and B_2 would be found as $20-5$, *i.e.*, 15 units, and in the second case it would be $25-7$, *i.e.*, 18 units. The error is in neither case significant.

The contrasts obtained in this way can be compared one with another in the manner indicated, but care must be taken not to place too fine an interpretation upon the figures obtained. For example, whilst 17 units (see above) can be stated with certainty to represent a much greater contrast than 2 units, it would be making too great an assumption to state that the first contrast (of 17 units) was $17/2=8{\cdot}5$ times as great as the second contrast (of 2 units). *All that can be expected of the method of assessing contrast is, that it will grade contrasts in order of magnitude more successfully than will a consideration of the physical brightness values alone.*

The consistency of the method can be estimated for one case in the following way. Dunbar* has obtained a relationship between the background brightness of a road surface and the brightness of an object seen against it, for the contrast between the object and its background to be just sufficient for a driver travelling at 30 m.p.h. to be certain of seeing the object. Since the criterion of contrast was the same throughout the range of background brightness

* *Trans. Illum. Eng. Soc. (Lond.)* 3, 187–195, Dec., 1938.

investigated, it would be expected that the contrast remained constant. Using his results and the data of fig. 6 we obtain Table 1.

The values of contrast obtained are reasonably constant.

In each separate problem it is necessary to correlate values of contrast with the appropriate limit cases. For example, in the case of street lighting problems, the data of Dunbar indicate that the minimum contrast between an object on the road and the background for a driver travelling at 30 m.p.h. to be certain of seeing the object is 5 units. Hence any contrast which under the relevant conditions is less than 5 units would be considered inadequate. In the same way the limit conditions for another visual task can be found by experiment, and contrast values obtained under test conditions compared with the limiting value.

In fig. 6 an arrow attached to each curve indicates the adaptation brightness for which the curve is drawn. This again shows (see fig. 3) that when the eye is adapted to moderately high brightnesses, the sensation corresponding to this adapting brightness is in the middle of the range of sensations; brightnesses lower than the adapting brightness cause correspondingly lower sensations. But when adapted for low brightnesses the adapting brightness itself causes little sensation, and brightnesses less than the adapting brightness cause practically none. All that can be seen are brightnesses above the adapting brightness, which cause a disproportionately great sensation.

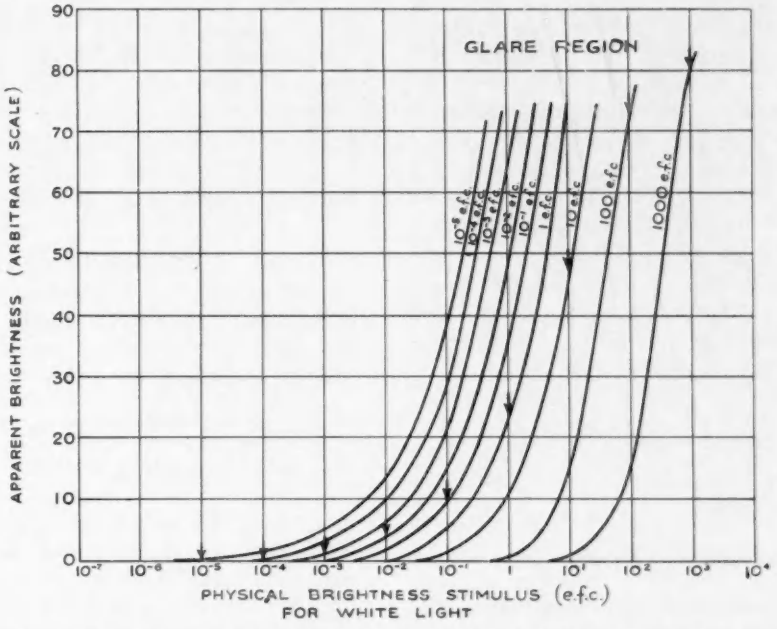


Fig. 6.—Averaged data showing relationship between apparent brightness and physical brightness for adaptation levels from 10^{-5} e.f.c. to 1,000 e.f.c.

TABLE 1.

Background Brightness (e.f.c.)	Object Brightness (e.f.c.) (from Dunbar)	Apparent Brightness		Contrast
		B'ground (from fig. 6)	Object	
1.0	0.7	23	18	5.0
0.3	0.19	17	12	5.0
0.1	0.06	11	6.5	4.5
0.03	0.01	7.3	3.3	4.0
0.01	0	5	0	5.0

This can be seen well in a familiar effect in the blackout. On dark starlit nights the white painted indications, white lines, etc., stand out surprisingly and appear almost luminous. The general level of brightness of these white lines will be about 6 or 7 times that of most objects present, and consequently they cause a sensation out of all proportion to that which they cause at higher levels. An example occurs below.

(*c*) *Incorrect measures of contrast.*

It has been pointed out in section 2 above, that contrast is a sensation, and that therefore it cannot be measured directly. The methods of section 6(*b*) serve to obtain a grading of contrasts in order of magnitude, but are not necessarily a strict measure of contrast.

This grading of contrast is, however, very much to be preferred to any method of assessing contrast on the basis of the stimulus alone. All previous methods have, of necessity in the absence of any

physical-brightness/apparent-brightness data, been measures of stimulus only.

Previous measures of contrast have been usually based on :—

(1) $\frac{B_1}{B_2}$ the brightness ratio, or

(2) $\frac{B_2-B_1}{B_2}$ *i.e.*, $\frac{\triangle B}{B}$, the relative brightness difference.

The function

(3) B_2-B_1 has been used by Stiles, who is always careful to describe it as a brightness difference ; it is sometimes misquoted as implying a "contrast."

The first measure, *i.e.*, the ratio of the contrasting brightnesses, will give an approximately correct grading of contrast when (i), the contrasting brightnesses are of the same order as the average brightness of the scene, and when (ii), this average brightness is between about 1 e.f.c. and 1,000 e.f.c. For conditions of low brightness, or when the contrasting brightnesses differ much from the adapting brightness, the errors involved may be very great. For example, the contrast between brightnesses of 10 e.f.c. and 100 e.f.c. at an adaptation level of 100 e.f.c. is, by (1), the same as between 0·001 e.f.c. and 0·01 e.f.c. at the same adaptation level, whereas in fact the latter contrast would be imperceptible, both being indistinguishable from black.

The second measure, $\frac{B_2-B_1}{B_2}$, will give a correct grading of contrast under the same conditions as the first measure. Both depend, of course, on a logarithmic relationship between physical brightness and apparent brightness. Such a relationship is known to hold over a limited range, as evidenced by the parts of the curves of fig. 6 which are approximately straight lines. Over this range, a given ratio of physical brightness always produces the same difference in apparent brightness.

The third measure, B_2-B_1, gives a rough grading of contrast *for a given adaptation level*, but even under constant adapting conditions the grading fails to discriminate contrasts correctly in the ranges where methods (1) and (3) above are reasonably close.

All such measures of contrast which do not take into account the adaptation level are, therefore, liable to considerable error, especially for all work concerned with low brightnesses.

For example, consider four juxtaposed brightnesses of 1, 10, 100 and 1,000 e.f.c. viewed at an adaptation level of 100 e.f.c. The contrast between each and its neighbour measured by the three methods above, and by the method of section 6(*b*) is given in Table 2.

The three contrasts A-B, B-C, and C-D are represented as equal by the measures B_2/B_1 and $\frac{B_2-B_1}{B_2}$; the measure B_2-B_1 would suggest that the contrast A-B : B-C : C-D : : 1 : 10 : 100. Visual inspection demonstrates immediately that none of these measures approaches the truth. The contrast by the apparent brightness method is assessed as A-B : B-C : C-D : : 14 : 57 : 30, which is in good accord with the visual impression.

TABLE 2.

	Brightness (e.f.c.)	B_2/B_1	$(B_2-B_1)/B_2$	B_2-B_1	Contrast by apparent brightness method
A	1	10	0.9	9	14
B	10	10	0.9	90	57
C	100	10	0.9	900	about 30
D	1000				

If now all the above brightnesses are reduced by 1/1000, so that four brightnesses of 0·001, 0·01, 0·1 and 1 e.f.c. are viewed at an adaptation level of 0·1 e.f.c. the "contrasts" between each brightness and its neighbour are as shown in Table 3.

Whilst the measures B_2/B_1 and $\frac{B_2-B_1}{B_2}$ still give the same values of "contrast" as before, B_2-B_1 suggests very much smaller contrasts, whilst the method of apparent brightness indicates that contrasts A-B and B-C are much less than previously, whilst contrast C-D is not reduced to such a degree. This is in accord with visual inspection.

If the scale of the experiment is now further reduced by 1 : 1000, to bring the brightnesses to within blackout range, the contrasts are as given in Table 4, the adaptation level now being 10^{-4} e.f.c. (corresponding to about weak moonlight).

TABLE 3.

	Brightness (e.f.c.)	B_2/B_1	$(B_2-B_1)/B_2$	B_2-B_1	Contrast by apparent brightness method
A	0.001	10	0.9	0.009	2
B	0.01	10	0.9	0.09	7
C	0.1	10	0.9	0.9	26
D	1				

The measures B_2/B_1 and $\frac{B_2-B_1}{B_2}$ still give the same contrast as before, which is obviously incorrect, whilst B_2-B_1 yields values of the wrong order, relative or absolute. The apparent brightness method, however, again accords with visual inspection in showing that the contrast A-B is now imperceptible, that B-C is just noticeable, and that C-D is relatively prominent.

These last two conditions again illustrate the effect already mentioned. The contrast C-D is so

much greater than the contrasts A-B and B-C that the brightness D appears to stand out prominently from the rest. At higher brightness levels, as shown by Table 2, however, the contrast C-D is actually *less* than the contrast B-C, because the brightnesses approach the glare region.

It is obviously legitimate (although not always useful) to express the *stimulus* conditions as

$$\frac{B_1}{B_2} \text{ or } \frac{\triangle B}{B}$$

or some other function of the brightnesses concerned; but it is desirable that such functions should be described accurately for what they are—as for example brightness ratio, or relative brightness difference, etc.—and not as "contrast" which they often are not. This will avoid much confusion which at present exists.

TABLE 4.

	Brightness (e.f.c.)	B_2/B_1	$(B_2-B_1)/B_2$	B_2-B_1	Contrast by apparent brightness method
A	0.000001 (10^{-6})	10	0.9	0.000009	0
B	0.00001 (10^{-5})	10	0.9	0.00009	1
C	0.0001 (10^{-4})	10	0.9	0.0009	2
D	0.001 (10^{-3})				

7. CONTRAST PHENOMENA IN PRACTICE.

It is interesting to note how many of our daily actions can be attributed to the effects of contrast. For example, sitting with one's back to the window at an interview has a twofold effect. The interviewer's face is in shadow, and the background to it—the window—is bright. The person being interviewed is therefore looking with eyes adapted to a high brightness at a not too well illuminated face, the brightness of which is near the lower end of his range, so that he can see little of its detail. His own face, however, receives higher illumination, and is is seen against a comparatively dark background; its brightness is near the middle of the range of the interviewer's eyes. All the details of the face and changes of expression are enhanced for one person and diminished for the other.

Many familiar devices for making things conspicuous operate by increasing contrasts. A cricket bowling screen; the white screen behind certain railway signals; the black, white and red of surveyors' station poles; the soldiers' elementary rule of avoiding the skyline, are all obvious examples. The use of white paint on kerbs and obstructions is another, the value of which has increased in the blackout for the reasons already explained. The importance of contrast in vision in lighted streets and by headlights needs no emphasis.

In industry it is necessary to avoid glare, either direct or reflected, because glare "makes seeing difficult"; this we had learned by experience. The curves of fig. 6 show some of its effects. Glare increases the adaptation level of the eye so that a given brightness difference produces a less sensation of contrast. (This is of course only one of several manifestations of glare.) On the other hand, for a given *ratio* of brightness between two objects the higher the absolute value the greater the contrast, within quite wide limits. Therefore the higher the illumination, in general, the better we see. This also can be seen from fig. 6.

The problem of enhancing contrasts for inspection, particularly for metal surfaces, is frequently that of applying the simple laws of reflection.[1]

Surprisingly enough these laws are often difficult to apply in practice, because the best answer may demand a costly double or triple inspection; but the lighting engineer cannot be blamed for this.

An interesting example of a fairly straightforward problem is that of razor-blade inspection. Among the many tests to which modern safety razor blades are subjected is the visual inspection shown in fig. 7.[2] A powerful light is directed on to the edges of the blades, which are arranged to lie parallel with one another, the blades being a little out of the vertical. A good edge, being very sharp, will reflect so little light to the inspector's eyes (obeying the laws of specular reflection) as to appear almost black. A flaw, on the other hand, will diffuse light over a considerable angle, and generally appears a contrasting gray. A tray of blades, having a large percentage of flaws, is shown in fig. 8 as it would appear to the inspector, with the general arrangement shown in fig. 7. To locate all the faults it is necessary to repeat the inspection with the tray turned through 180° and the blades reset so that they assume the same angle to the light as before, the reverse side now facing the inspector. The background to the lighting system must be dark.

Perhaps the most remarkable application of contrasts and their behaviour is the portrayal in a picture having two dimensions of objects having three dimensions. Several aspects of this have already been discussed. Presumably it is possible because in practice the eye perceives depth only by inference and not directly; it distinguishes directly only differences of brightness and colour over various directions in space.[3] If these can be simulated in a picture or photograph, the visual sense is reminded of the brightness and colour distribution of the original, with a vividness depending upon the

1. W. R. STEVENS and S. ANDERSON; *Trans. Illum. Eng. Soc.* (London), Vol. III., No. 1, January, 1938.

2. Figs. 7 and 8 are reproduced by kind permission of the Ever Ready Razor Co., Limited.

3. At close ranges accommodation, convergence and the stereoscopic sense also assist in the inferring and estimation of depth, but their effect is surprisingly small.

faithfulness with which the brightness and colour distribution is reproduced.

It has been emphasised that geometry is perhaps

Fig. 7.—General arrangement for inspection of razor blades.

the most important single item in recognition. It is not only the geometry of the outline of an object, but also that of the detail within the boundary. For example, a change of shadows within a face can alter the expression completely. This is the main device on which an actor's make-up depends. Until comparatively recently, an actor would treat his face rather as a blank canvas on which most of the shadows had to be painted, for the "flat" lighting usually adopted cast very few shadows. This is altered somewhat to-day, for modern stage lighting technique is able to introduce shadow contrasts, and therefore interest, in a way which was impossible with earlier and cruder apparatus.

Contrast may create interest, but one of the most important war-time applications of contrast effects is camouflage, which is of course, the craft of "not attracting interest." When a building is camouflaged it cannot disappear completely, but it may become so inconspicuous in relation to its surroundings, that an observer does not notice it, although if he does notice it, it may seem as easy to see as before the camouflage was applied.

The illuminating engineer may divide the processes of camouflage into three :—

(i). Camouflage by reducing contrasts.

(ii). Camouflage by introducing artificial heavy contrasts which conflict with the true geometry of the object (disruptive camouflage).

(iii). Camouflage by introducing artificial contrasts similar to those associated with a different object (imitative camouflage).

The most familiar example of type (i) is the khaki army uniform, which "blends well" into a variety of landscapes (*i.e.*, produces small contrasts with its usual backgrounds). The camouflage consists of a more or less uniform colour, matching as nearly as possible that of the surroundings. The arctic fox, which in winter changes its coat to pure white, is a good example from nature ; the camel, matching the colour of the desert sand, is another. From modern military practice, apart from khaki uniforms, there are the "duck-egg blue" colour on the undersides of R.A.F. day aircraft to match the sky, and the gray of warships, which is varied according to the part of the world in which they are operating.

If this type of camouflage is applied to fixed positions the natural surroundings may be altered in appearance according to the time of year. Therefore, if a building is well camouflaged in summer with greens, it may require russet browns in autumn and grays and whites in winter. The colours which made it inconspicuous in one season may enhance the contrasts in another. So this type of camouflage needs attention to keep pace with the changing appearance of the countryside.

Disruptive camouflage aims at breaking up the *outline* of the object by large areas of colour, arranged so that they produce geometrical patterns which conflict with the true geometry of the building and its shadows. The effect is to produce a building really quite easy to see, but hard to recognise.

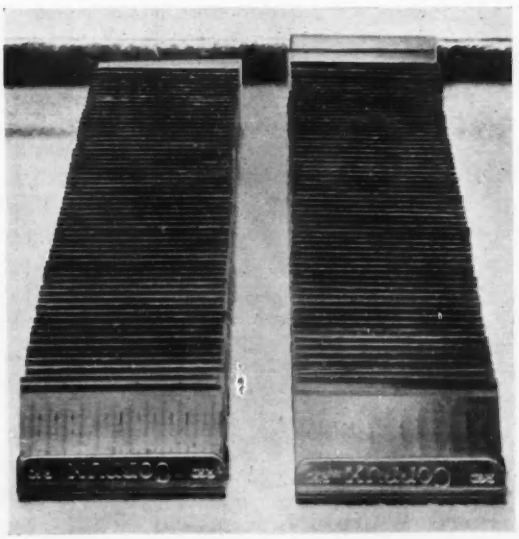


Fig. 8.—Faulty razor blades as seen by the inspector.

Examples are so common that it is unnecessary to go into details. In nature it occurs in animals such as the tiger, the zebra and the leopard. The sizes of the patches must be comparable with that of the

outlines which they are intended to confuse. If they are too small they will themselves become inconspicuous before the building does when viewed from a distance.

By the third method of camouflage (imitative camouflage) the building is made uninteresting by introducing contrasts associated with something else. A factory becomes a row of houses; a machine-gun post a hayrick. The idea behind the method is obvious; its application often complex. The method also occurs in nature, for example in the "stick insect," which looks like a twig. It has the advantage of not needing to be changed to match changing surroundings as often as the other types of camouflage.

8. CONCLUSION.

This paper has described some of the first principles of sight, and has applied them to a few simple cases. They have necessarily been greatly simplified; and in simplifying them there is a danger of supposing that the matter has been nearly exhausted, and that little remains to learn.

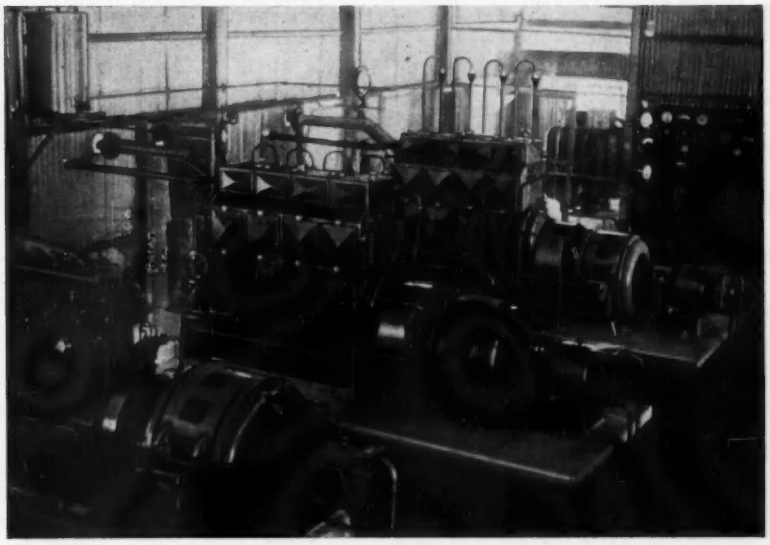
The physics alone is extraordinarily complex. Each of the myriads of points which make any scene is reflecting radiation with an individual distribution in all directions in space; the energy is radiated on a whole spectrum of frequencies, the spectral distribution varying from one point to another, and even for any one point varying often with the direction of reflection. Then the whole of this complex of radiation varies with time.

A minute fraction of it enters the eyes by two apertures, a few millimetres in diameter; and in a fraction of a second it is analysed and classified, as to its directions, its intensities and its spectral distributions, and this vast information is passed on so quickly that we can follow easily its variation with time. It is a system of analysis and classification surpassing anything else in Nature.

But even the physics is less wonderful than the interpretation of the information. From it we deduce continuously form, motion, size, distance, texture, material, position, of many objects at once; age, character, expression, behaviour; almost all that we know of everything that we cannot touch. In the end, most of the material for our imagination, art, and beauty, come in the same way.

Our best experiments are crude by comparison; we have touched but the fringe of the matter, and we can only stand in awe of so marvellous a mechanism. Yet without some understanding of seeing we cannot begin to serve our sight, which is the proper business of illuminating engineering.

Equipment of a Power House in the East.



Engine driven alternators in a power house supplying a large mining undertaking in the East. This station provides power for motors driving air compressors, pumps, haulages and milling machinery; and its entire electrical equipment, including the overhead transmission lines, was supplied by The General Electric Co., Ltd. In addition, a considerable portion of the mining plant was manufactured by Fraser & Chalmers.

Phosphorescence and its Applications.

By J. N. BOWTELL, B.Sc., and E. E. MILES.

G.E.C. Research Laboratories.

INTRODUCTION.

WHEN a suitable stimulating agency, such as ultra-violet radiation, is allowed to fall on certain substances, they are seen to glow brightly with a characteristic colour. In other words, these substances are transforming ultra-violet energy into visible light. This phenomenon is known as photo-luminescence.

The property of exhibiting luminescence is not confined to any particular class of substance. Actually nearly all substances fluoresce to some extent. However, the luminescent materials which are discussed here are those specially prepared in the form of crystalline powders to give the maximum possible luminescent effects. The manufacture of these materials presents many difficult problems owing to the extreme care required to exclude all traces of accidental impurities. This becomes a serious difficulty owing to the fact that the powders have to be furnaced at a high temperature during manufacture.

The present article is chiefly concerned with a discussion of phosphorescent materials, that is to say luminescent materials which have the property of storing some of the incident energy and re-emitting it as visible light after the exciting source has been removed, as distinct from fluorescent materials which emit light only while irradiated.

Until a few years ago phosphorescent effects, whilst of considerable academic interest, had found but little practical application. In the past few years, however, these effects have been extensively studied, and methods of producing phosphorescent materials of high quality have been developed. War-time conditions have, moreover, brought about many new and important applications of these materials.

HISTORICAL.

Phosphorescent effects attracted little or no scientific attention until the beginning of the 17th century when a certain Vincenzo Cascariola of Bologna accidentally produced, whilst dabbling in alchemy, a form of barium sulphide which gave an orange red glow in the dark. This became known as Bologna stone. In the following two centuries a number of substances which phosphoresced with a characteristic colour after exposure to light were discovered. At this time each newly discovered material was named after its discoverer, and the word "phosphorus" was applied as a general term. Thus one reads of "Homberg's Phosphorus" (calcium chloride), "John Canton's Phosphorus" (calcium sulphide), and "Balduin's Phosphorus" (calcium nitrate).

During this period Robert Boyle succeeded in isolating a substance which glowed in the dark without previous exposure to light. For a time this material was classed with other known phosphorescent materials and was named "Boyle's Phosphorus." This substance was actually the element phosphorus which, as is now well known, exhibits a greenish yellow glow whilst undergoing oxidation in moist air.

In 1867, the Frenchman, Sidot, produced a brilliantly green phosphorescent zinc sulphide compound by heating specimens of the mineral sphalerite. This discovery was of great importance since this luminous substance was the forerunner of the highly efficient zinc and zinc-cadmium sulphide materials which are being applied in many fields at the present time.

At about this time it was first recognised that luminescence was produced only when the substance was in a crystalline state, and furthermore was markedly influenced by the degree of purity of the sample. High purity was essential for the best results but the presence of traces of certain heavy metals was also necessary. In compounds produced by earlier workers the necessary impurity occurred fortuitously in the starting substances used.

Towards the end of the last century an improved form of phosphorescent calcium sulphide, which glows with a blue colour, was prepared by heating together calcium and sulphur with a trace of bismuth. After mixing with a suitable varnish, the preparation, which was marketed under the name of Balmain's Paint, was applied to match-boxes and the like to betray their whereabouts in the dark. This possibly represents the first commercial application of phosphorescent materials. The sulphides of the two other alkaline earth metals, barium and strontium, can be produced in a phosphorescent form. Strontium sulphide, which usually exhibits a green-blue or yellow-green afterglow of particularly long duration, is one of the most important compounds now used to produce luminous signs for A.R.P. and other war-time uses.

As regards fundamental investigations on the subject of luminescence, the first contribution of outstanding importance was made in 1852 by Stokes who enunciated the law which bears his name. This states that the wavelength of the exciting radiation is always less than that of the fluorescent or phosphorescent light produced. Furthermore, by filtering off the visible radiation from sunlight and

Fig. 1.—Spectra of luminescent phosphors.
(a) Mercury spectrum for comparison
(b) Zinc sulphide activated with copper
(c) Strontium sulphide activated with bismuth

leaving only the invisible ultra-violet radiation, Stokes was able to distinguish clearly between fluorescent and phosphorescent effects. At about the same time Becquerel was studying the decay of phosphorescence in different substances, and devised a simple instrument for this purpose. In this, the material was coated on a drum which could be rotated at known speeds. The phosphorescent coating is excited from one side of the drum and viewed on the other. In this way it was possible to observe the phosphorescence at a definite time after excitation.

At the present time decay curves (as they are called) are often obtained by direct visual photometry when the durations of afterglow are a matter of minutes or hours. For the measurement of phosphorescent effects of short duration, photoelectric photometers coupled to automatic recording devices are employed.

GENERAL.

Since the beginning of the present century the art of preparing inorganic luminescent materials has made enormous strides, and the present-day materials are by comparison extremely efficient. All inorganic luminescent materials of the type under discussion exhibit phosphorescence to a greater or less degree and are now referred to as phosphors. Of these, only a few of the luminescent sulphides are sufficiently phosphorescent for practical purposes.

All the inorganic phosphors are characterised in having a crystalline structure and their luminescent properties largely depend on the presence in the crystal lattice of minute traces of certain metallic impurities or activators. The spectrum of the light emitted by these materials consists of one or more broad bands whose position depends primarily on the parent substance or matrix, as it is called, and the activator. For example zinc sulphide activated by copper gives a green band, whilst the same substance activated by silver gives a blue band. If, however, whilst retaining copper as activator, the matrix is changed by combining with the zinc sulphide an appreciable quantity of cadmium sulphide, which has a similar crystalline structure, the emission band can be shifted towards longer wavelengths. In fact luminescence of any colour from green to red may be obtained by progressively increasing the quantity of cadmium and using a suitable activator. Fig. 1 shows the spectra of zinc sulphide activated with copper and of strontium sulphide activated by bismuth.

EFFECT OF WAVELENGTH ON EXCITATION EFFICIENCY.

A phosphor when undergoing excitation does not convert all wavelengths of the incident radiation into visible light with equal efficiency. In fig. 2(*a*), the curve marked A represents the emission band of a typical phosphor. The shape and position of this curve is practically unaffected by changing the wavelength of the exciting light. If only fluorescence is considered, and the efficiency of different exciting wavelengths is measured, curve B in fig. 2(*b*) is obtained. Fluorescence will not be excited by wavelengths lying outside this curve, but will be most efficiently excited by monochromatic radiation of wavelength corresponding to its peak. Similarly curve C in fig. 2(b) represents the form of the excitation curve for phosphorescence. It will be seen that the two curves B and C are not coincident, but light whose wavelength lies in the region where

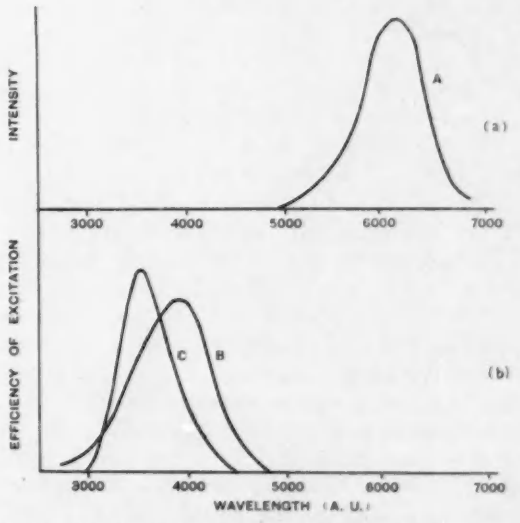


Fig. 2.—Forms of emission band and excitation curve.
(a) Emission band.
(b) Curve B : excitation curve for fluorescence.
Curve C : excitation curve for phosphorescence.

they overlap will excite both fluorescence and phosphorescence.[1]

EFFECT OF INFRA-RED ON PHOSPHORESCENCE.

The effect of long wave infra-red radiation on phosphorescence is interesting. Gently warming a glowing phosphor causes a very marked brightening or, as it is called, acceleration of the light emission. The heated region finally becomes darker than its surroundings since the amount of stored energy in the

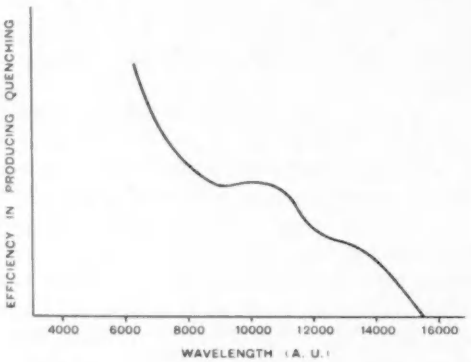


Fig. 3.—Efficiency of different wavelengths in producing quenching of phosphorescence.

heated region has been reduced. It has been shown that provided infra-red radiation of sufficiently long wavelength is employed, the total amount of stored energy emitted as phosphorescence is not reduced by this accelerating process. Gently heating a material which has been kept in the dark for a considerable time and in which the intensity of phosphorescence has decayed to a value below the threshold of visibility, produces a burst of phosphorescence, showing that even after a long period the phosphor retains some stored energy. The acceleration effect has been used to measure the total integrated amount of stored energy emitted as phosphorescence in a given material.

Roughly speaking, it is found that if all the stored energy were released from a layer of powder in a period of half-a-second, the brightness would be about that of an ordinary candle flame.*

In contrast to the effect of long wavelength infra-red, irradiation of phosphors by short wavelength infra-red radiation and red light produces a quenching effect on phosphorescence, in which part of the stored energy is released as heat, and the total amount emitted as light is consequently reduced. The effectiveness in producing this quenching is dependent upon the wavelength of the infra-red radiation. Fig. 3 shows a curve of efficiency in producing quenching plotted against wavelength for a typical phosphor. It will be seen that orange and red radiation is far more effective in this respect than the longer wavelengths. The quenching effect is sometimes accompanied by the acceleration effect mentioned above, whereby the phosphor will show a short intense burst of phosphorescence followed by a fairly quick extinction.

* In actual fact, the total stored energy is of the order of 500,000 ergs per gram in the case of a good quality strontium sulphide phosphor.

MECHANISM OF LUMINESCENCE.

In the foregoing paragraphs mention has been made of some of the more important properties of inorganic phosphors. An attempt to link these together in terms of the modern theory of luminescence is beyond the scope of this article. An elementary explanation based on present knowledge of the processes involved is, however, given.

To study the mechanism of luminescence, it is usual to picture in the crystal a series of broad electron energy bands associated with each atom or group of atoms in the matrix. These bands may be imagined as regions surrounding a group of atoms in the matrix much in the same manner as the electron orbits surround a single atom in a gas. For convenience, they are usually represented diagrammatically by horizontal bands as shown in fig. 4. The electrons linked with the atom groups can only have energies corresponding to one or other of these bands which are separated by so-called forbidden regions as shown in fig. 4(*a*).

In the normal unexcited state the lower band is supposed to be filled with electrons whilst the upper band is empty. When the energy in the absorbed radiation is sufficient to raise an electron from the lower to the upper band, fluorescence would be caused if the excited electron returned to the lower band. Present-day theories, however, indicate that the electron does not normally return directly to the band it has just vacated. Fluorescence is actually made possible by the effect of the metal activator atom in introducing new energy levels near to the lower band as shown in fig. 4(*b*). In these circumstances, when an electron is raised from the lower to the upper band, the vacated space is immediately filled by an electron from the so-called impurity level, and fluorescence is caused by the excited electron in the upper band returning to the vacant space now created in the impurity level. The energy difference between the upper band and impurity level is less than that between upper and lower band, hence the fluorescent light will be of lower frequency or longer wavelength than the exciting light, which is in accordance with Stoke's Law.

As an explanation of phosphorescence, further vacant energy levels are postulated near to the upper band as shown in fig. 4(*c*). These trapping levels, as they are called, are thought to be partly due to the activator and partly due to strains and other abnormalities set up in the matrix during the heat treatment necessary to form the crystal. Excited electrons

from the upper band may be trapped in these levels from which they can only be released again to the upper band by thermal excitation. These electrons may return to the lower band at a considerably later stage, thus producing phosphorescence.[2, 3]

This simplified picture, whilst illustrating in a general way the mechanism of all the activated inorganic phosphors, is particularly useful in explaining the more important properties of the sulphide materials. The dependence of the colour of the luminescence radiation and the position in the spectrum of the excitation bands on the particular matrix and activator is seen to be due to the different spacing of the energy bands and impurity levels in the different materials. Again the simple theory outlined above would lead one to expect the colour of the emitted light to be largely independent of the wavelength of the exciting radiation, and in practice this is found to be true for most phosphors. The picture described above not only gives an explanation of phosphorescence, but indicates the reason for the varying shapes of decay curves encountered in different phosphors. When a phosphor

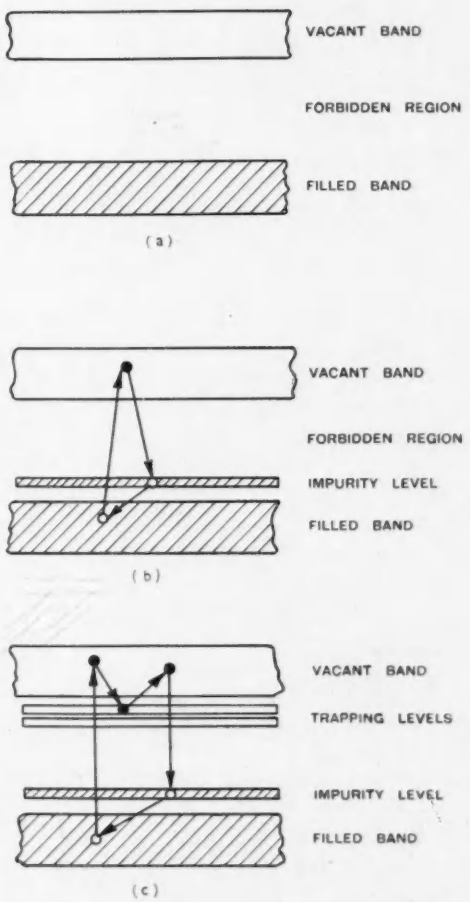


Fig. 4.—Diagrammatic repesentation of the energy bands in a phosphor.

is exposed to suitable exciting radiation, the trapping levels become filled with electrons. The fluorescence radiation observed is mainly due to electrons returning from trapping levels which retain electrons for a very short time, that is those situated close to the upper band. These electrons return to their normal levels within a very small fraction of a second, so that when the exciting radiation is removed, a rapid initial drop in the emitted light takes place. Electrons situated in trapping levels further removed from the upper band will return at a later stage, thus producing the normal long term phosphorescence. If the trapping levels are mostly situated close to the vacant band, the decay curve will take the shape of that shown by curve A in fig. 5. That is to say there will be a relatively intense burst of phosphorescence which rapidly decreases to a very low value. On the other hand, when the trapping levels are mainly situated at some distance from the vacant band, the initial fall in phosphorescence intensity may be very steep, but the tail of the curve may continue for some hours at a comparatively high level, as shown by curve B in fig. 5. It can thus be seen that the actual shape of the decay curve in any given phosphor, will depend on the detailed distribution of the trapping levels.

Since the trapping levels are supposed to be partly due to localised strains arising in the crystal lattice during its formation, it is to be expected that heat treatment, mechanical crushing and other processes through which a phosphor passes during manufacture may greatly influence its duration of afterglow, and in actual practice this is found to be the case. The existence of trapping levels also readily explains the acceleration of light emission in a glowing phosphor by warming, mentioned earlier; since it is postulated that electrons already in the trapping levels are freed to the upper band by thermal excitation, it follows that any increase in temperature of the phosphor will accelerate this process of freeing, and thus induce a brightening of the afterglow.

Further evidence for the existence of trapping levels is provided by the effect of low temperature on phosphorescence. On cooling a glowing phosphor in liquid air, the phosphorescence energy is "frozen in" and remains so until the phosphor is warmed up, when the emission continues. The effect may be explained by the fact that, at low temperatures, there is insufficient thermal energy in the crystal to "trip over" the electrons from the trapping levels to the upper band.

BRIGHTNESS DECAY CURVES.

The most important phosphorescent materials from the practical point of view, are the well known zinc and zinc-cadmium sulphide and the alkaline earth sulphide phosphors. The zinc and zinc-cadmium sulphide series which are produced com-

mercially in a wide variety of colours, are highly fluorescent, but for the most part possess relatively short afterglow. It is possible, however, to prepare one or two members of this series in a highly phosphorescent form.

The alkaline earth sulphide phosphors as a class exhibit a much longer phosphorescence than the zinc and zinc-cadmium sulphides, but their initial brightness is considerably lower. Of these, the most suitable for practical purposes is strontium sulphide activated by bismuth which gives a blue-green phosphorescence. Fig. 6 shows the decay characteristics for typical samples of the two types after irradiation by a low intensity source of ultra-violet. If the lowest useful brightness be taken as 0·0001 equivalent foot-candles, it will be seen that the decay time to this level for the zinc and strontium sulphides is very approximately one hour and seven hours respectively. Excitation by a more intense source would give a somewhat longer decay time. The curve also shows that in the first 3 or 4 minutes of the decay period, the brightness of the zinc sulphide is much greater than that of the strontium sulphide, but after this period, the latter remains the brighter.

Brightness decay curves of the type shown in fig. 6 cannot be satisfactorily expressed by a single formula over their whole range. The useful region of the curve which occurs after the first few seconds is of the form[4] :—

$$I_t = \frac{I_0}{(1 + \alpha t)^2}$$

I_0 = Brightness of phosphor at time 0.
I_t = Brightness of phosphor t secs. after irradiation has ceased.
α = a constant.

The difference in decay characteristic of the two materials discussed above will largely govern the manner in which they are used in practice. On account of their high fluorescent brightness, the phosphorescent zinc sulphides are particularly suitable in locations where, under normal conditions, they may be irradiated by a convenient source of ultra-violet light, but where, in the event of failure of the electrical supply, a useful brightness must be maintained by phosphorescence for a reasonable period. Such situations include the marking of important electrical and other control points. Owing to the low fluorescent brightness of phosphorescent strontium sulphide, separate means for exciting these materials would not normally be employed. They are, however, well excited by tungsten lamps and by daylight, and their long afterglow characteristics render them suitable for the production of various outdoor signs, direction and obstruction indicators, etc., for civil defence and allied purposes.

DARK-ADAPTATION AND APPARENT BRIGHTNESS.

A phosphorescent material having been excited, for example, by the normal tungsten lighting of a room will appear, when viewed in the dark by an observer whose eyes had been adapted to the same lighting, to glow at approximately constant brightness for a considerable time. This is due to the fact that the decay in brightness of the phosphor is largely compensated by the increasing sensitivity of the observer's eyes.

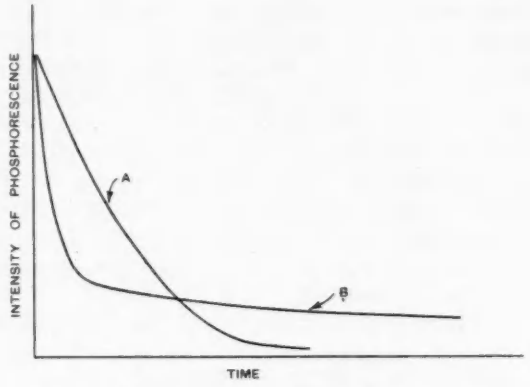


Fig. 5.—Forms of decay curves of phosphors.

At the relatively low levels of brightness obtaining with phosphorescent materials, their apparent brightness is largely dependent on the condition of the observer's eyes, and to a great extent, this fact governs the usefulness of these materials in any given location.

To an average eye in the fully dark-adapted state, a brightness of between 0·0003 and 0·0002 equivalent foot-candles can be fairly readily perceived in completely dark surroundings provided that the luminous area is of reasonable size.

However, to an eye adapted to the level of brightness of a well lighted room, a brightness one hundred times greater, *i.e.*, 0·03 e.f.c. would be barely perceptible under the same conditions, and only after a lapse of several minutes in the dark would the eye be sufficiently rested to perceive a brightness of 0·0003 e.f.c. Thus, a phosphorescent material, particularly one which had not been excited or "charged" for a considerable period, whilst giving a useful indication to a dark-adapted person might be invisible to an observer coming from a well lighted room.

The effectiveness of phosphorescent materials used out-of-doors depends largely on the darkness of the night. On very dark nights, strontium sulphide phosphors of good quality excited during the hours of daylight will glow at a useful level of brightness throughout a long night. On moonlight nights, on the other hand, the material may give no better indication than would ordinary white paint.

SOURCES OF LONG WAVE ULTRA-VIOLET RADIATION.

Lamps of the kind generally known as "black lamps" are the most suitable sources of long wave ultra-violet radiation for the excitation of phosphorescent materials. In addition to the 80 watt and 125 watt "Osira" black glass lamps, for example, the 6 watt "Osglim" argon-filled lamp may be used. The former are similar in general construction and operation to the well known 80 watt and 125 watt "Osira" h.p.m.v. lamps, except that the outer bulbs are made of a special black or nickel glass which absorbs the visible radiation from the discharge, whilst freely transmitting the long wave U.V. The black glass bulb completely absorbs all harmful short wave radiation, and may safely be used without any special precautions having to be taken to protect the eyes. This class of lamp is one of the most efficient sources of long wave U.V. available, emitting as it does a considerable fraction of its total radiated energy at 3650A.

The "Osglim" argon-filled lamps have a special gas filling giving a negative glow discharge rich in near ultra-violet radiation. There is very little visible radiation from the discharge so that the use of a black glass filter is for many purposes unnecessary. These small lamps, although much less efficient than "Osira" black glass lamps as generators of U.V., nevertheless have many uses where only small amounts of U.V. are required, as for example, in luminescent "Exit" or similar small box signs.

As will be seen from curve B in fig. 2(*b*), the peak of the excitation efficiency curve of sulphide phosphors occurs near 3650A. Thus the lamps described above are particularly suitable for exciting these powders.

Tungsten lamps used in conjunction with special black glass filters serve as convenient, if very inefficient, sources of violet and long wave U.V. The relative efficiencies of the three sources of near U.V. mentioned above may be seen from Table 1 which gives the brightness of a sample of green luminescent zinc sulphide when irradiated at a distance of 3ft.

TABLE 1.

Relative Performance of Various Sources of Long Wave U.V. Radiation.

Source (without reflector)	Brightness of sample of green luminescent zinc sulphide at 3ft. from source	Approx. relative U.V. efficiencies (Black glass lamp = 100)
125 watt "Osira" black glass lamp	5.0 e.f.c.	100
230 volt, 150 watt Osram tungsten lamp and black glass filter ..	0.07	1.2
6 watt "Osglim" argon lamp (no filter) ..	0.03	12

The 1940 Lighting (Restrictions) Order requires that all signs used for traffic guidance and A.R.P. purposes should be legible at a distance of 100 ft. and inconspicuous at 250 ft. in the absence of mist, fog, rain and moonlight.

In using these sources of U.V. for exciting luminescent materials outdoors, care should be taken that the values of brightness obtained do not exceed those necessary to fulfil this requirement. The various maximum brightnesses permitted are given in British Standard Specification ARP/32. Another point to be borne in mind is that no U.V. radiation in a direction above the horizontal is permitted. (See British Standard Specification ARP/18).

STABILITY OF PHOSPHORESCENT MATERIALS.

Both the zinc and zinc-cadmium sulphide and the alkaline earth sulphide phosphors deteriorate to a greater or less degree, with consequent loss of luminescence, if left unprotected in the atmosphere. In the case of the zinc and zinc-cadmium sulphide

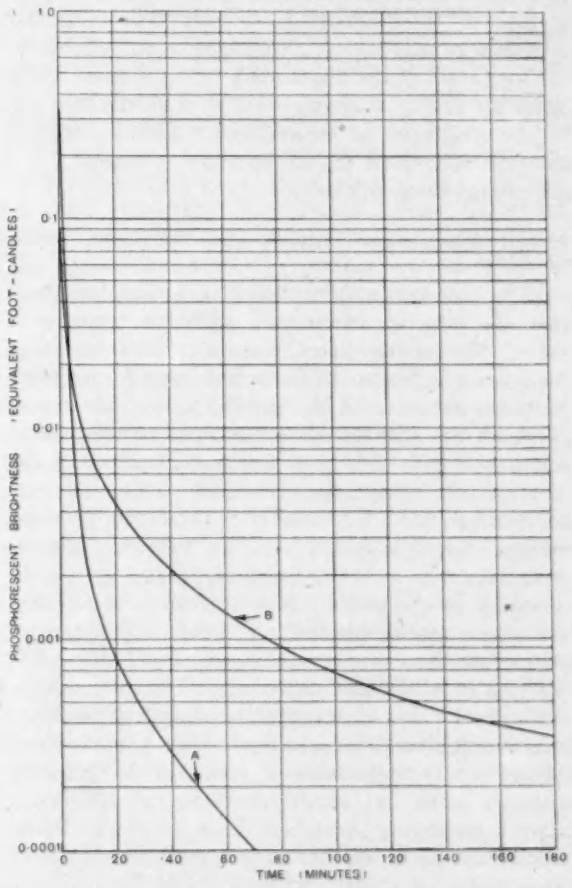


Fig. 6.—Curves of decay of phosphorescent brightness.
(a) Green luminescent zinc sulphide.
(b) Green-blue luminescent strontium sulphide.

powders, exposure to sunlight in an atmosphere containing moisture causes more or less rapid blackening. It appears that the combined action of radiation shorter than 4500A and moisture produce chemical reduction of the surface of the crystals. For given conditions of exposure to light and moisture, the rate of blackening is affected in large measure by the detailed treatment of the phosphor during manufacture. Irradiation of the powder in vacuo by blue light or near-ultra-violet radiation of high intensity over long periods causes little or no blackening.

In certain circumstances, the complete exclusion of moisture is, however, not essential for the prevention of blackening. Thus an aqueous solution of potassium silicate has a good preservative action when used as a binding medium for these powders.[5] In fact, for application by spraying, such a binder may be employed with good results provided the product is not subjected to outdoor weathering conditions.

Deterioration of the alkaline earth sulphides can proceed in total darkness if moisture is present, and is due to their tendency to hydrolyse with the evolution of sulphuretted hydrogen, accompanied by the formation of the metal oxide or carbonate. Here again the rate of decomposition is markedly affected by the conditions of manufacture. In this case the complete exclusion of all moisture is essential for preserving these materials.

PHOSPHORESCENT PAINTS AND PAINTING PROCESSES.

The zinc and strontium sulphide phosphorescent materials may be conveniently utilised in the form of paints by mixing them with a suitable binding medium or vehicle. In the actual mixing, care must be taken not to crush the minute individual crystals of which the powder is composed, since this would seriously impair the luminescence. The formulation of a suitable vehicle for the powder is a matter of some difficulty since a number of conflicting requirements must be fulfilled. The vehicle consists essentially of a transparent film-forming plastic dissolved in a suitable solvent to which is added a small amount of plasticizing agent. The correct choice of these three components and the proportions in which they are used govern the qualities of the final paint with regard to ease of application, film forming properties, and adhesion. In addition, owing to the deterioration which these powders undergo when left unprotected, the vehicle must exert a preserving action on them under the worst conditions likely to be encountered in practice. They must, of course, transmit freely both visible and near-ultra-violet radiation.

The development of suitable vehicles possessing the properties outlined above was undertaken by I.C.I. (Paints), Ltd. in collaboration with the G.E.C. Research Laboratories. The media finally produced have a good preservative action on the powder, and have excellent working properties from the application point of view. To obtain maximum durability, it is necessary to apply these media in accordance with a carefully worked out painting schedule, details of which are given in the Appendix.

In the application of luminescent paints to various surfaces, certain precautions have to be taken to ensure (*a*) maximum brightness, and (*b*) good durability.

(*a*) Apart from the intrinsic efficiency of the phosphor, the important factors affecting brightness are the coating density (or weight per unit area) of the luminescent material in the paint layer, and the reflection factor of the surface to which it is applied. As the coating density is increased, there is a rapid rise in the brightness at first, but this becomes progressively less for equivalent increases in weight. Fig. 7 gives the relationship between fluorescent brightness and coating density for a zinc sulphide paint applied to a white background. It can be seen that a value of about 150 gms. per sq. metre is a reasonable concentration. A similar curve for phosphorescence is shown in fig. 8 which gives the time of decay to a brightness of 0·0002 e.f.c.

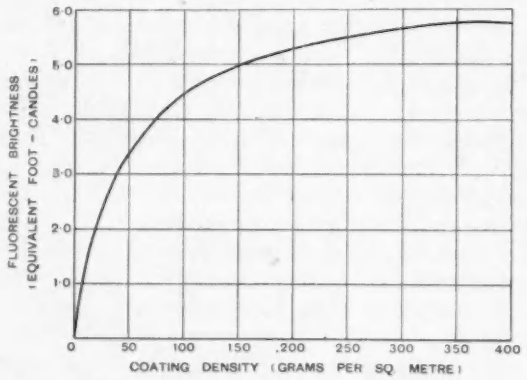


Fig. 7.—Relation between fluorescent brightness and coating density for a green luminescent zinc sulphide phosphor irradiated by a 125 watt "Osira" black glass lamp at a distance of 3 feet.

for the same material plotted against coating density. Here a reasonable concentration is 175 gms. per sq. metre. For strontium sulphide paint, the concentration for phosphorescence under the same conditions is 150 gms. per sq. metre, and the time of decay to 0·0002 e.f.c. is 5 hours.

With regard to the effect of background colour, it can be stated that with zinc sulphide material applied to a concentration of 150 gms. per sq. metre, the fluorescent brightness is increased by approximately 25 per cent on changing from a black to a white background. It is therefore advantageous to

apply phosphorescent paints to a white background.[6]

(*b*) For any given binding medium, the ability of phosphorescent paints to withstand weathering is largely dependent upon the detailed method of application. To ensure the maximum durability on both indoor and outdoor locations, the exact painting schedule recommended by the manufacturer should be used, and great care should be exercised in obtaining uniform coatings in the actual brushing or spraying work.

Luminescent paints may be applied to wood or metal. After the normal preparation of the surface, an undercoat of white paint is applied to ensure maximum brightness. The paint employed for this should be lead-free otherwise premature blackening of the phosphor may occur. The phosphorescent paint proper is then applied to the appropriate thickness followed by a protecting coat of the unpigmented clear medium. Each coat must, of course, be thoroughly dry before the application of the next. By dilution with a suitable thinner, application by spray-gun may be carried out and the necessary thickness of coating built up by applying several thin coatings. Owing to the greater protection necessary for the alkaline earth sulphides, the coating schedules are usually somewhat more elaborate than for the zinc sulphide type.

Detailed painting schedules for both types of paint are given in the Appendix.

It is not advisable to apply phosphorescent paints directly to existing walls, wood and metal work, etc., but to prepare independently painted metal panels which may be fixed up where required.

All paints, whether luminescent or not, are

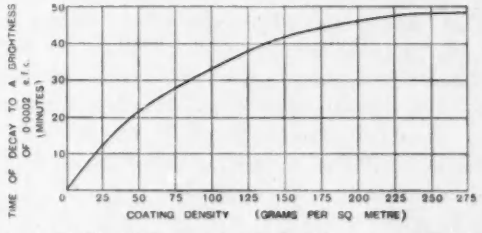


Fig. 8.—Relation between the time of decay to a brightness of 0.0002 e.f.c. and coating density for a green luminescent zinc sulphide phosphor.

liable to deteriorate in service. Apart from the nature of the paint medium and method of application, the rate of deterioration is, of course, particularly affected by such variable factors as location and weather conditions. It is possible to give only a rough indication of the likely reduction in efficiency of luminescent paints under average conditions. In the case of the media mentioned above, applied by their respective painting schedules, little or no reduction in brightness should be expected on indoor locations over a period of a year or so. On outdoor positions a useful life of from three months to one year may be achieved with alkaline earth sulphide paints according to the circumstances. Luminescent zinc sulphide paints are considerably more resistant to weathering than the alkaline earth type and their outdoor performance is correspondingly better.

For permanent use in very exposed positions, a method has been developed in the G.E.C. Research Laboratories whereby the alkaline earth sulphide paints can be almost completely protected from weathering. This consists essentially of forming a sandwich of the paint between two pieces of sheet glass, the edges of which are carefully sealed with clear medium. A symbol may then be stencilled in black on to one side of the glass. Sandwiches up to about eight inches square can readily be produced in this way. As an extension of this scheme larger signs may be made by sandblasting the necessary letters and symbols in plate glass to a depth of about 2 mms. The hollows thus formed are filled with phosphorescent paint and allowed to dry. The whole of the back surface is then coated with a black paint consisting of the special clear medium mixed with a suitable pigment. The finished plate is then backed with a sheet metal which acts both as further protection and as a means of supporting the plate. Signs for A.R.P. purposes constructed in this manner should be made with reinforced plate glass.

The process of silk screening is widely used for the rapid production of signs, posters, etc., and can be used with the luminescent paints described above, provided certain precautions are taken. The apparatus is similar in principle to the ordinary hand-operated silk screen duplicator. The silk screen used readily allows the passage of the phosphorescent paint. The formulation of the paint media described above is modified to reduce their drying rates. Care must be taken to ensure a sufficiently thick coating of phosphor by this method, and two or three coats of the phosphorescent paint should be applied to build up the recommended thickness of coating. The over-coats of clear medium are usually applied by spray-gun when the phosphorescent paint is dry. All the precautions taken when applying by brush or spraying methods should also be observed in this process to ensure good durability. With regard to the squeegee used for forcing the paint through the screen, it is advisable for this to be made from some material other than rubber, since the particular solvents employed in phosphorescent paint media tend to attack rubber.

PHOSPHORESCENT SOLID SYNTHETIC RESINS AND VITREOUS ENAMEL.

Phosphorescent sulphides have been successfully incorporated in solid synthetic moulding plastics and resins, and moulded into a variety of

shapes such as switch covers, nameplates and other articles. The durability of such phosphorescent plastics depends on the type of resin used, but generally speaking it is about the same as painting according to the best recommended technique.

Zinc sulphide phosphors may also be mixed with a suitable enamel frit for the production of a phosphorescent enamel, and excellent samples of phosphorescent enamels have been produced. In this case the durability under outdoor conditions is comparable with that of the best paints. Alkaline earth phosphors have not been successfully incorporated in vitreous enamels.

PRACTICAL APPLICATIONS.

Under the conditions brought about by the present lighting restrictions, phosphorescent materials may be used to advantage in the solution of a number of wartime lighting problems, and particularly in rendering luminous many types of sign used for Air Raid Precautions and similar purposes.

Considering in this connection the use of the irradiated zinc sulphide types, advantage may be taken of the fact that a large number of signs, arrows, etc., can be excited by the one U.V. source. Thus a single 125 watt "Osira" black lamp suitably placed will excite phosphorescent signs to a useful brightness over a radius of 50 yards.

Typical instances of this method of use are the irradiation of the entrances to a group of shelters, and of control equipment, instrument dials in power houses, etc. As previously mentioned, in the event of a supply failure phosphorescent zinc sulphides will maintain a useful brightness for about half an hour. For dispersed signs it may be necessary to employ an individual ultra-violet source for each sign. For such a purpose the 6 watt "Osglim" argon-filled lamp is particularly suitable. This lamp may also be successfully used to irradiate a number of closely grouped articles such as instrument dials, etc.

The long afterglow of the alkaline-earth sulphide materials enables these to be used where a suitable electricity supply is not available, or where it is not considered economical to install extra supply points. The alkaline-earth sulphides, as already mentioned, are, in general, much less bright in use than the irradiated zinc sulphides. In certain circumstances, such as during a long night preceded by a very dull day, they may also give disappointing results owing to the fact that insufficient energy has been absorbed during the hours of daylight. For this reason it is inadvisable to use these materials for "key" situations unless provision exists for exciting them independently.

Experiments have shown that the alkaline earth sulphide phosphors may be successfully used for a variety of purposes. Amongst these may be mentioned all types of signs for A.R.P. purposes outlining entrances to shelters, first-aid posts, decontamination centres, etc.; helmet lettering for A.R.P. personnel and arm-bands for pedestrians; luminous "S.P." and other symbols such as indicating plaques for the houses of firewatchers. They may also be used for route numbers on trams and 'buses, street names and house numbers, for aids to movement on station foot-bridges, steps, etc., and temporary indicators for street and other obstructions.

They may, however, prove to be unsatisfactory for certain purposes such as the direct marking of roadways, kerbs and pedestrian crossings where they are unlikely to withstand the normal traffic wear.

The highly fluorescent, as distinct from phosphorescent, zinc and zinc-cadmium sulphide materials have been very successfully applied in poster advertising, decorative shop window lighting, stage effects, etc.; a source of ultra-violet radiation is, of course, installed to excite the materials.

Novel and attractive effects may be obtained in many cases by combining fluorescent and phosphorescent materials in the same scheme which is then irradiated on a predetermined cycle with "off" periods at regularly spaced intervals. By this means also, attention can be drawn to the more important features of advertisements, shop window displays and the like. A somewhat similar application to poster advertising in carriages of underground trains has been found successful in practice. Here the salient feature of the advertisement is "picked out" in phosphorescent paint and remains visible after the carriage lights are extinguished when the train comes above ground.

Phosphorescent articles such as electric light switches may be used to advantage where work proceeds mainly in the dark. Thus in photographic dark rooms it is convenient to make a number of switches and other controls phosphorescent. Provided reasonable care is used in the disposition of these, trouble due to fogging of even the most sensitive photographic materials is not likely to arise.

Phosphorescence has been employed in a system of infra-red telegraphy in which the receiver comprises a moving luminous tape and the transmitter an apparatus for the production of a parallel beam of infra-red radiation which can be interrupted for the sending of code messages. The beam is brought to a focus on the tape which shows as the message is received, a number of dark dots and dashes due to the "quenching" effect of the infra-red radiation.

Use can also be made of the quenching effect of infra-red to produce rapidly a positive picture from a photographic negative. The negative to be "printed" is held against a sheet coated with phosphorescent material which has been previously excited. A powerful beam of short wavelength infra-red radi-

ation is projected for a few seconds on to the negative. A "positive" is revealed on the phosphorescent sheet since the less dense regions of the negative in transmitting more infra-red give rise to darker regions on the corresponding area of the sheet, and vice versa. The process is of course carried out in a darkened room.

The property of phosphorescence is made use of in luminescent discharge lamps to reduce the slight stroboscopic flicker caused by the extinction of the discharge at the current minimum in each half-cycle. Although the powders used in these lamps are for the most part only slightly phosphorescent, this helps to bridge the out period and reduces the flicker.[7]

SELF-LUMINOUS MATERIALS.

These consist essentially of a fluorescent material of the zinc sulphide class intimately mixed with a very small amount of a radium or radio-active salt. As is well known these radio-active salts emit radiation continuously which excites the zinc sulphide crystals in much the same way as does ultra-violet radiation. In earlier self-luminous materials the radio-active substance employed was a radium salt which possessed the disadvantage that the intensity of fluorescence varied over a wide range during its life, the brightness reaching a maximum in about 20 days and then gradually declining to a constant low level after about 300 days.[8] In recent years, by the use of meso-thorium in place of radium, it is possible to obtain practically constant luminosity over a period of more than two years, after which there is a slow decline.

These self-luminous materials are usually mixed with a suitable varnish for use in paint form. Owing to their high cost (approximately 100 times that of non-radio-active phosphors) they can only be used to cover very small areas where cost is not a primary consideration.

APPENDIX.

PAINTING SCHEDULES.

Phosphorescent Zinc Sulphide Paints for Brush Application.*

(1). Apply one or two brushed coats of recommended lead-free white paint. Allow to dry thoroughly.

(2). Apply two brushed coats of phosphorescent paint. Allow each coat 2 hours to dry. (The phosphorescent paint is made by carefully stirring 1 kg. of powder into 1 litre of the appropriate binding medium.)

(3). Apply one brushed coat of the clear medium. Allow 2 hours to dry.

Phosphorescent Strontium Sulphide Paints for Brush Application.*

(1). Apply one or two brushed coats of recommended lead-free white paint. Allow to dry thoroughly.

(2). Apply two or three brushed coats of phosphorescent paint. Allow each coat 1 hour to dry. (The phosphorescent paint is made by carefully stirring 0·3 kg. of powder into 1 litre of the appropriate binding medium.)

(3). Apply one (preferably two for outdoor locations) brushed coats of the clear medium. Allow each coat 1 hour to dry.

Phosphorescent Strontium Sulphide Paint for Spray-gun Application (Stoving type).

(1). Apply one sprayed coat of recommended lead-free white paint. Stove for one hour at 110°C.

(2). Apply two sprayed coats of phosphorescent paint. Allow to dry off for 15 minutes between each coat, and stove the double coat for ½ hour at 110° C. (The phosphorescent paint is made by carefully stirring 0·4 kg. of powder into one litre of the appropriate binding medium. About ⅓ litre of recommended thinner is then added.)

(3). Apply two sprayed coats of the clear medium suitably thinned for spraying, allowing 15 minutes between coats, stoving the double coat for ½ hour at 110°C.

(4). Apply one sprayed coat of the clear medium as used for (3), and stove for one hour at 110°C.

* For applying by spray-gun dilute with recommended thinner.

REFERENCES.

1. J. W. Ryde. Luminescent Materials and their Application to Light Sources. *Trans. Ill. Eng. Soc.* (London) 3, No. 8 (1938).
2. J. T. Randall. Luminescence and its Applications. *Jour. Roy. Soc. Arts*, 85. p. 354 (1937).
3. R. P. Johnson. Luminescence of Sulphide and Silicate Phosphors. *Jour. Opt. Soc. Am.*, 29, p. 387 (1939).
4. F. Seitz. *Trans Faraday Society*, 35, p. 75 (1938).
5. British Patent Spec., No. 502919.
6. W. E. Harper, Margaret B. Robinson and J. N. Bowtell. The Photometric Properties of Some Luminescent Materials. *Trans. Ill. Eng. Soc. (London)* 5, p. 57 (1940).
7. L. Levy and D. W. West. Modern Application of Luminescent Substances. *Chemistry and Industry*, 62, p. 457 (1939).
8. C. C. Paterson, J. W. T. Walsh and W. F. Higgins. An Investigation of Radium Luminous Compound. *Phys. Soc. Proc.*, 29, p. 215 (1917).

Television Film Transmitters using Apertured Scanning Discs.

PART I.*

By D. C. ESPLEY, D.Eng., M.I.E.E., and D. O. WALTER,
G.E.C. Research Laboratories.

Various methods are outlined which can be used for the production of interlaced signals by apertured-disc film scanners. A novel arrangement using an intermittent-magnification optical system is described, and details are given of a complete transmitter according to this method. The continuously moving film is illuminated by a split optical system to remove picture flicker, and a device is included for the exact compensation of film shrinkage.

This paper is reprinted from the Journal of the Institution of Electrical Engineers, Part III., June, 1941.

(1) INTRODUCTION.

THE scanning methods and apparatus described in this paper are the result of an investigation into forms of transmitters for 405 lines most suitable for standard interlaced picture-signal generation. The quality assessment of modern television receivers demands that the test picture signal should be above suspicion as regards resolution, maintenance of black level, freedom from shading effects, noise level, accuracy of interlacing, and linearity of scanning. It was decided to adopt the apertured disc as the primary scanning member in a film transmitter arranged to produce a wave form substantially the same as that used for the London television service.[1]

Apertured discs for sequential scanning have been commonly used from the beginning of television,[2] but interlacing methods by discs have hitherto received attention mainly in Germany where developments have reached a commercial stage. The authors were able to gain experience with one of these transmitters for 240 lines, and the scanning system (trapezium projection) is described below. Although early constructional work on the 405-line transmitter was based on this technique, a novel method of scanning (intermittent magnification) was developed and incorporated in the final design.

As far as the authors are aware no comprehensive account of a complete disc transmitter has appeared in any journal. In particular, the information available on disc construction is very meagre, and development had to proceed from first principles. The location of the centre of each scanning aperture is a problem in itself, but when it is shown that the area of each aperture has to be about 2×10^{-6} sq. in. it is easy to account for the characteristic dark lines which appear in the pictures from many disc scanners owing to complete or partial blocking of the apertures by foreign matter. A special technique had to be developed for disc cleaning, and it was found that such cleaning could be done initially and would not require repetition after installation of the disc in a vacuum chamber.

It can be shown that a scanner according to the described method is not limited to 405 lines. At present the maximum resolution is set by the ultimate strength and density of disc materials rather than by signal/noise ratio.

An arc lamp is used as the light source and it is considered that as cinema practice is so well established such an illuminant presents no operational difficulties.

A complete transmitter installation constructed in accordance with the principles outlined in this paper, but without illumination and film-shrinkage compensation, was first demonstrated to television experts in December, 1937. The construction of a disc scanner presents severe difficulties, but it is expected that the apparatus can maintain quite constant picture-signal characteristics over an almost unlimited life.

(2) METHODS OF INTERLACING BY DISC.

The broad idea of scanning by means of a rotating disc is well known and was originated by Nipkow in 1884.[2] Scanning by a disc of the type then proposed would produce a simple scan of sequential or non-interlaced form. Many of the problems associated with disc scanning do not arise until interlaced signals are required. The basic difficulties of interlacing can best be illustrated by considering first of all the simplest arrangement of scanning by means of apertures located along the length of a continuous and moving band. Although the method is only of small practical interest it has the greatest freedom from inherent distortions, and a comparison indicates the nature of the optical problems which arise in the

* NOTE.—The numerical references in the text are to the Bibliography which, together with the Appendices, will be published with Part II of this article.

preferable mechanical arrangement of the plane circular disc.

In fig. 1, four possible scanning methods are shown by making use of the relative motion between an illuminated picture film and a continuous strip which carries repeated groups of, for example, 7 holes each. For simplicity it is assumed that the planes of the strip and the film are coincident, but of course in a practical example the film would be replaced by an optical image in the plane of the strip. Cases (*a*) and (*b*) produce a sequentially scanned picture, made up of 7 horizontal lines dis-

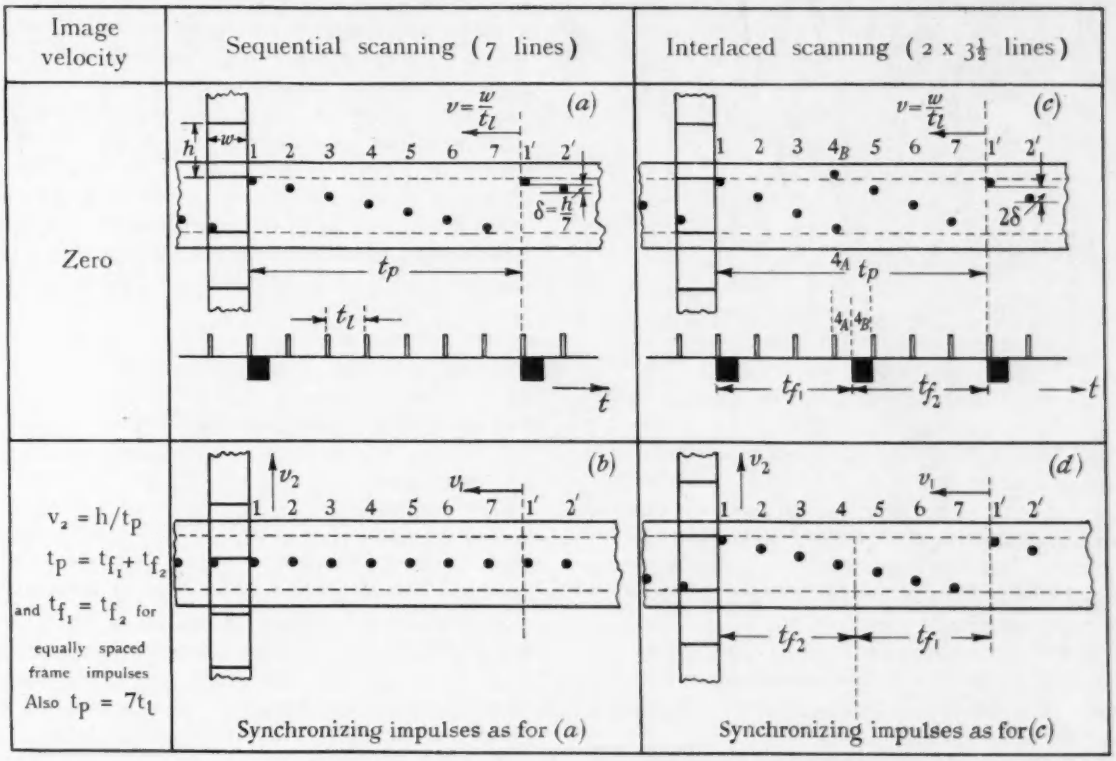


Fig. 1.—Scanning of image by strip with rectilinear motion.

posed continuously from top to bottom of the picture, if the hole centres are spaced longitudinally by the picture width w, and move forward with a velocity $v_1 = w/t_l$, in which t_l is the line scan time. Case (*a*) differs from case (*b*) in that the former is arranged with a group of staggered holes displaced by a unit transverse increment $\delta = h/7$ to co-operate with a film which is stationary for the picture repetition period t_p and is then jerked forward by h to allow the next picture to cover the scanning field.

Intermittent motion of the film by the Maltese Cross mechanism of the standard cinematograph projector would be satisfactory except for the fact that about 30 per cent of the time is lost in film transit, and no television system could afford the increase of frequency band required to transmit the picture information during the remaining stationary period of 70 per cent.

Case (*b*) demonstrates the favoured modern system in which the film moves with a constant velocity, and it is then possible to dispose the holes in a straight line parallel to the direction of strip motion. By compounding the velocities of film and strip it will be seen that the holes trace substantially the same paths across the film in both cases, except for a small inclination $\gamma = \text{arc tan}\left(\frac{ht_l}{wt_p}\right)$ when the film is moved continuously. This error is negligible when the number of lines is large ($\gamma \sim 6{\cdot}4$ minutes for 405 lines), but may be corrected by inclining the axis of the film by an angle γ.

We are here concerned with the production of interlaced signals of the "odd line" or "non-integral" type. By this is meant the signal form now standardized by the B.B.C. for the Alexandra Palace transmitter,[1] and in which there is an odd number of lines per picture and an even number of interlaced frames. Strip-scanning methods for such signals are shown in fig. 1(*c*) and 1(*d*), and the film motions correspond to (*a*) and (*b*) respectively. In case (*c*) the hole centres are displaced by a transverse increment 2δ, except at the boundaries of a frame scan. It will be seen that holes 1, 2 and 3 scan non-contiguous strips spaced by 2δ, and it is arranged

that hole 4_A scans the first half of line 4 to finish one frame, and the scan is continued by hole 4_B at a point one picture-height above 4_A. The scan is completed by holes 5, 6 and 7 in a total time t_p. The change-over from 4_A to 4_B can be determined by an optical shutter, but this is not essential in a high-definition system as the synchronizing process occupies a short period in this region, and a few holes near the jump point could even be omitted from the strip without any effect on the picture.

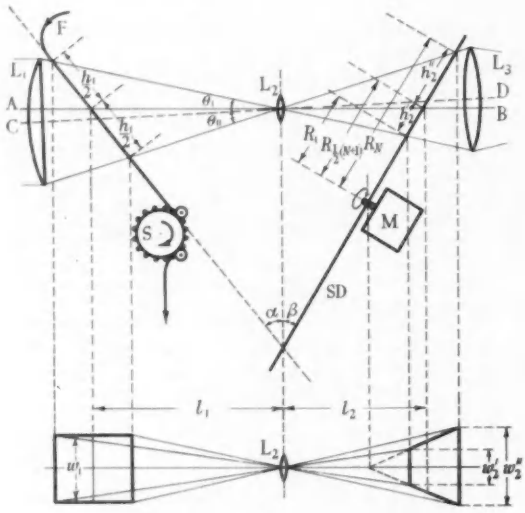


Fig. 2.—Trapezium projection system.

Case (*d*) is of the most interest as it represents the basic method adopted in the present transmitter. The holes are located exactly as for (*a*), but the film moves with a constant velocity h/t_p. In the diagram, holes 1, 2, 3 and 4 (half line) are about to scan the picture for the second frame before the picture leaves the scanning field.

At the receiver the luminous scanning spot must move in exactly the same manner as do the successive scanning holes relative to the film picture mask, and in this way the electrical signal, which is variable in time, is translated into an immobilized picture. It is most convenient and economic to arrange that the receiver spot moves along the line with constant velocity and amplitude, and this is standard practice. As the picture is built up of two superimposed frames it is essential that the scale of the two frames should be the same, or at least variable only slightly over the area in a manner exactly the same for both frames. In this requirement resides the inherent problem of interlacing by a scanning disc. The problem does not exist in case (*d*), as all the holes move with equal velocities and in the same direction. The picture information is then always scanned at the same rate, and thus the ratio of hole velocity to receiver-spot velocity is always a constant.

(2.1) Trapezoidal Projection.

The scanning disc for interlaced signals, but without geometrical corrections, is essentially the strip of fig. 1(*d*) bent into a complete plane annular ring with hole 4 at the mean radius. The holes then lie on a spiral of Archimedes, of pitch h and with a radial increment between consecutive hole centres of $h/7$. This step introduces two distortions:

(*a*) As the holes are on a spiral the hole velocity is continuously variable from hole to hole and hence from beginning to end of the scan, and it is not feasible to vary the angular velocity of the disc over one revolution. The holes must pass a radial line at equal time-intervals owing to receiver time-base features already mentioned. This variation of line-scanning velocity is tolerable if the mean radius of the spiral is very much greater than the pitch h, and if simple sequential scanning were used according to fig. 1(*a*), the only result would be to produce a slightly trapezoidal or tapered picture at the receiver. This effect gives rise to a further distortion of interlaced signals which is altogether inadmissible, as the average velocity of the holes for one frame is different from the average for the second frame, with the result that the two frames at the receiver are to different scales and thus fail to register.

(*b*) The holes do not move in straight lines, so that the scanning field is arcuate in shape. This is a fundamental property of all scanning discs for sequential and interlaced signals, but the effect is almost imperceptible as the holes in both frames are subject to the same error, and there is no departure from registration although the picture information in both frames is bowed slightly in a direction perpendicular to the line. A method of correction[21] is available, but as it produces a curved image from a rectangular object by the use of lens aberrations, it imposes at present too many restrictions on the design of the optical system. In what follows this distortion will be neglected.

Some experience had been obtained, on other transmitting equipment made and supplied by Radio Loewe A.G., for 241 lines, of a very elegant method,[3] due to Schlesinger, of reducing the main distortion (*a*) by projecting on to the scanning disc an image of trapezoidal shape to fit the scanning field. Work on the present equipment was started along these lines and would have provided a theoretically exact solution of the geometrical problem, but it was decided to develop a new method with a view to ultimate comparison between the two systems. The trapezium method is illustrated in fig. 2 for the case of a lens or pinhole of very small aperture, and it is somewhat analogous to the technique employed in photographic enlarging for the correction of the keystone effect in negatives which have been taken with a tilted camera.

An elevation and plan of the system are shown

in the diagram, but for reasons of clarity the trapezium effect is very much exaggerated. The film F is moved by a sprocket S at constant velocity in a plane inclined by an angle $(\alpha + \beta)$ to the plane of the disc SD which is driven by motor M. The rectangular film picture has dimensions $h_1 \times w_1$ and is imaged by the lens L_2 in the plane of the disc.

$$h_2'' = \frac{h_1 l_1 m\sqrt{(\cos^2\alpha + m^2\sin^2\alpha)}}{2l_1 - h_1(1+m)\sin\alpha} \quad . \quad . \quad (4)$$

in which the nominal optical magnification m is given by:

$$m = \frac{l_2}{l_1} = \frac{\tan\beta}{\tan\alpha} \quad . \quad . \quad (5)$$

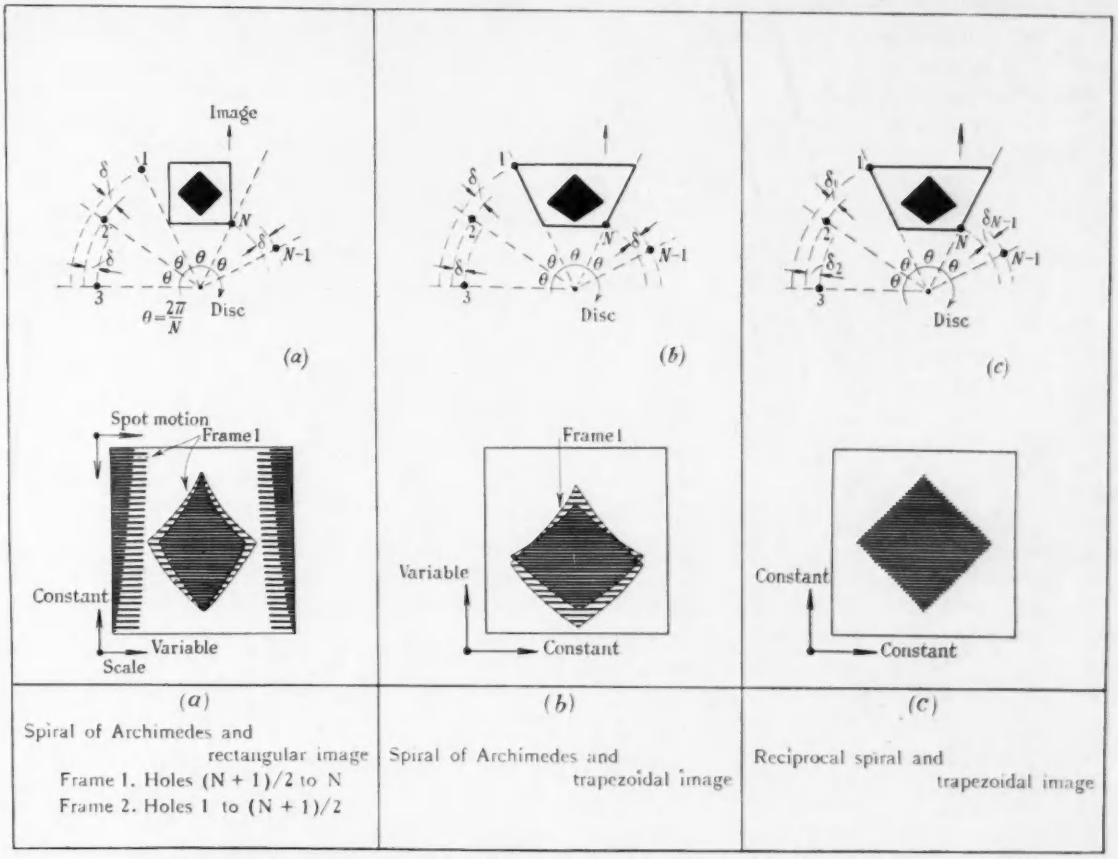


Fig. 3.—Interlaced scanning with spiral discs: effects of scale errors.

The image is a trapezium of height $(h_2' + h_2'')$ and of width varying from w_2' to w_2'' at disc radii of R_1 and R_N respectively. The film is illuminated by a source which is imaged by lens L_1 within the pupil of lens L_2. Light passing through the disc apertures is collected by lens L_3 to form a second image of the source on the cathode of a photocell where the signal current is originated. The axis AB of the object-image system is slightly inclined to the axis CD of the physical rays.

The geometry of the system is specified by the following equations:

$$\tan\theta' = \frac{h_1\cos\alpha}{2l_1 + h_1\sin\alpha} \quad . \quad . \quad (1)$$

$$\tan\theta'' = \frac{h_1\cos\alpha}{2l_1 - h_1\sin\alpha} \quad . \quad . \quad (2)$$

$$h_2' = \frac{h_1 l_1 m\sqrt{(\cos^2\alpha + m^2\sin^2\alpha)}}{2l_1 + h_1(1+m)\sin\alpha} \quad . \quad . \quad (3)$$

The trapezium sides w_2' and w_2'' have the values

$$w_2' = \frac{2w_1 l_1 m}{2l_1 + h_1(1+m)\sin\alpha} \quad . \quad . \quad (6)$$

$$w_2'' = \frac{2w_1 l_1 m}{2l_1 - h_1(1+m)\sin\alpha} \quad . \quad . \quad (7)$$

Thus

$$\frac{w_2'}{w_2''} = \frac{h_2'}{h_2''} \quad . \quad . \quad . \quad (8)$$

It can be shown that if the middle hole [number $\frac{1}{2}(N+1)$, and N is odd] is at a disc radius $r_{\frac{1}{2}(N+1)}$ the value of α is given by

$$\sin^2\alpha = \frac{l_1^2 m^2}{r^2_{\frac{1}{2}(N+1)}(1+m)^2 + l_1^2 m^2(1-m)^2} \quad . \quad (9)$$

The operation of the correction processes will be understood by reference to fig. 3, in which various stages of the correction are shown. The simple arrangement (*a*) is without correction as a rectangular

image is scanned by an Archimedian spiral of holes. Non-uniform hole velocity introduces at the receiver a variation of scale which is a minimum at the top of the picture and a maximum at the bottom. The vertical scale is correct. The diamond pattern is reproduced without registration of the two frames and sloping straight lines in the image appear curved as shown. As the image is first scanned by

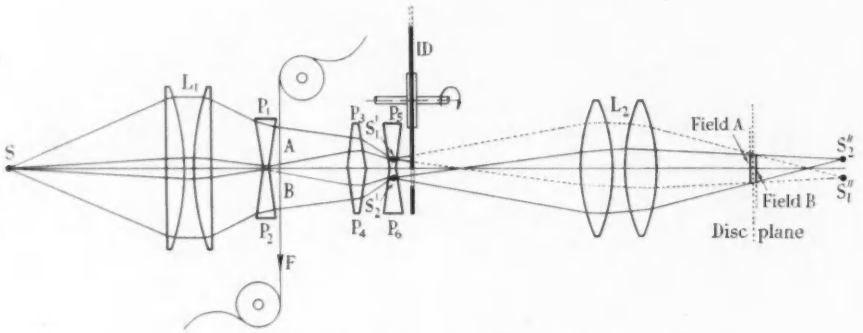


Fig. 4.—Optical system for overlapped scanning.

the inner holes the horizontal scale of frame (1) is always greater than that of frame (2), as it is assumed that the receiver spot velocity is a constant in both vertical and horizontal directions. The sides of the picture can never fit a rectangular mask.

The first correction of trapezium projection is applied in arrangement (*b*), and the received picture fits the mask. The horizontal scale is now correct, but the vertical scale increases from top to bottom owing to the variable velocity of the image as it moves across the scanning field.

A perfect scan is obtained if this effect of image-velocity variation is compensated by introducing a variable radial increment for the spiral of holes.[4] It is clear from inspection of the distorted images of arrangements (*b*) and (*c*) that the radial increment must be a minimum for the inside holes and be continuously variable over the scanning field.

The spiral must be of a special form and is best specified in terms of the radius of the mean hole, $r_{\frac{1}{2}(N+1)}$, and the trapezium coefficient ε, which is defined by :

$$2\varepsilon = \frac{w_2' - w_2'}{w_2'} \sim \frac{r_1 - r_N}{r_N} \quad \cdot \quad \cdot \quad (10)$$

The radius of any hole of number n is

$$r_n = r_{\frac{1}{2}(N+1)}\left[\frac{N(1+\varepsilon)}{N+\varepsilon(2n-1)}\right] \quad \cdot \quad (11)$$

and the variable radial increment is

$$\delta_n = r_n - r_{n+1} = \frac{r_n^2}{r_{\frac{1}{2}(N+1)}}\left[\frac{2\varepsilon}{N(1+\varepsilon)}\right]\left[1 - \frac{2\varepsilon}{N+\varepsilon(2n+1)}\right] \quad \cdot \quad \cdot \quad \cdot \quad (12)$$

The term $2\varepsilon/[N + \varepsilon(2n + 1)]$ is in practical cases very much less than unity (about 10^{-4}) and may be neglected. The increment may be taken as being very nearly proportional to the square of the hole radius, and the holes, if spaced by equal angular intervals, lie on the reciprocal spiral $r\theta$ = constant. A simple relation exists between the radii of holes 1, $\frac{1}{2}(N + 1)$ and N, viz.

$$\frac{2}{r_{\frac{1}{2}(N+1)}} = \frac{1}{r_1} + \frac{1}{r_N} \quad \cdot \quad \cdot \quad (13)$$

There are some disadvantages in the method, but these are more concerned with difficulties of construction than with theoretical limitations.

(2.2) Overlapped Scanning.

Several arrangements for continuously moving

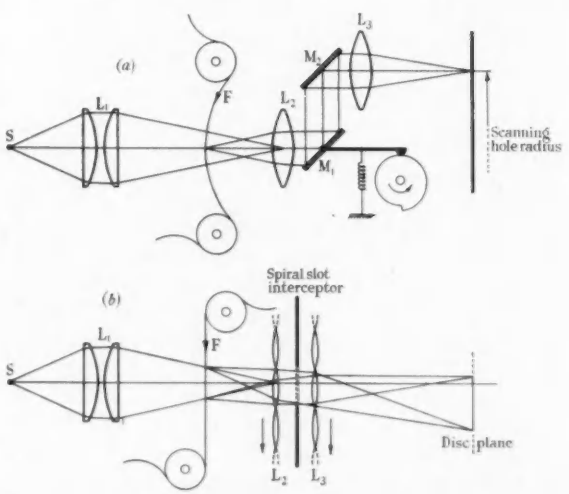


Fig. 5.—Optical arrangements for image displacement or immobilization.

film have been proposed in which interlaced signals are produced by scanning the first frame by a spiral array of holes, and then introducing a discontinuous optical shift of the image so that the same array scans the second frame. Such methods are most

suitable for even-line interlacing in which the frame impulses are staggered. One example[5] is shown in fig. 4. A double-light source, or single-light source S with two field prisms P_1 and P_2, illuminates the two halves of the film gate A and B via condenser L_1. Two images S_1' and S_2' would be produced on either side of the axis, but prism combinations P_3 P_4 and P_5 P_6 refract the beams so that S_1' and S_2' are both displaced towards the axis and lie within prisms P_5 and P_6 respectively. The distance between the prism combinations is adjustable and is arranged to produce equal and opposite film-image displacements, as shown in the diagram, of amplitudes just sufficient to overlap the appropriate image fields (A) and (B). These image fields are separately available by the simple shutter action of an interceptor disc ID, and complete interlaced scanning of N lines takes place by a single array of $\frac{1}{2}N$ holes within the common image field.

(2.3) Optical Compensators.

As long as the required relative motion is maintained between the disc holes and the film image, the

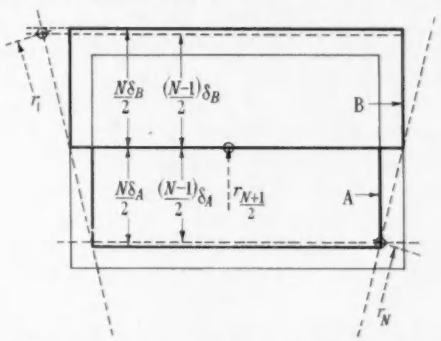


Fig. 6.—The geometry of intermittent magnification.

absolute velocity of either is immaterial. In one arrangement the radial velocity of the scanning apertures is zero,[6] and the axis of the optical system is turned at such a speed that the complete film frame image is moved at constant velocity past the holes in 1/50 sec., after which a discontinuous jump restores the image to the extreme position before the scanning of the second frame. The principle of operation is illustrated in fig. 5(*a*), in which it will be seen that a cam-operated mirror M_1 produces the necessary image shift. In this way it is possible to locate all holes on a circle, but odd-line interlacing is only available if the disc rotates a non-integral number of times per frame. This restriction imposes great difficulties in disc manufacture, as holes on opposite sides of the disc are closely related in the received picture, and effects of the nature of eccentricity produce first-order errors. It is also necessary to provide very accurately controlled discontinuous motion of the mirror M_1. The method could be so modified that complete immobilization of the image is introduced, and the disc is then provided with two Archimedian spirals so displaced from each other radially that interlaced scanning of the stationary image is achieved.

Unfortunately, if the two spirals are disposed over the same arc of the disc the radial displacement between them is only about one hole diameter, unless again a disc with non-integral turns is used.

Image displacement or immobilization can be achieved by the use of several optical systems, and mention should be made of the Mechau arrangement[7] and the moving projection lens system, of which an example has been embodied in a Farnsworth film scanner[8] (image dissector, and not scanning disc).

Fig. 7.—Interlaced scanning by intermittent magnification.

The latter system is shown in fig. 5(*b*) in which L_2 and L_3 are the moving components of the projection lens. These lenses are mounted on rotating discs, and a wide-slot interceptor disc obscures all lenses until required in the optical path.

(2.4) Intermittent Magnification (I.M. System).

The new form of scanner developed for the complete transmitter installation described in this paper uses an optical system with a magnification which is variable in time,[9] instead of in space as in the trapezoidal projection method. The basis of the method can be explained by reference to fig. 3(*a*). The absence of correction has introduced a distorted picture in which the vertical scale is the same for both frames, but the horizontal scale is variable from the top to the bottom in a continuous manner. The first step then is to increase the horizontal scale of the second frame until exact picture registration is attained. This can be done by increasing the magnification of the projection system for the period of the second scan, but one result of this would be that the normal two-dimensional changes of image size would vary the scale uniformly in the vertical direction. Fortunately this incidental effect can be compensated by adopting a different spiral pitch for the second scan. The geometry is shown in fig. 6, and it will be seen that if the geometrical relationship between the smaller image (A) and the scanning field for holes $\frac{1}{2}(N+1)$ to N is exactly the same as that between image (B) and the scanning field for holes 1 to $\frac{1}{2}(N+1)$, then the same picture scale appears for the two scans. The trapezium coefficient ε is defined as :—

$$\varepsilon = \frac{r_1 - r_{\frac{1}{2}(N+1)}}{r_1} \quad . \quad . \quad . \quad (14)$$

The constant radial increments δ_A and δ_B are related by :—

$$\frac{\delta_B}{\delta_A} = \frac{r_1}{r_{\frac{1}{2}(N+1)}} = \frac{r_{\frac{1}{2}(N+2)}}{r_N} \quad . \quad . \quad (15)$$

and it follows that $\delta_A = \dfrac{2\varepsilon r_{\frac{1}{2}(N+1)}}{N-1}$ . . (16)

In designing a disc of this form it is convenient to fix $r_{\frac{1}{2}(N+1)}$ and r_1, as $(r_1 - r_{\frac{1}{2}(N+1)}) = h_1 m_B(N-1)/(2N)$, in which m_B is the maximum magnification of the optical system. The magnification ratio $m_B/m_A = \delta_B/\delta_A$.

separate problem. It was decided at the outset that no reciprocating parts could be tolerated, and that the cycle of operations must be repeated exactly every picture to avoid the brightness flicker introduced by the differences between a number of optical paths. It is not necessary that the magnification change should take place instantaneously, as a time equal to the duration of the frame black margin is available for this purpose.

The principle is shown in fig. 8. In the first diagram a lens in plane P produces an image of height y_2 from an object y_1. If $l_1 = l_2$ then it is always possible to move the lens to a plane P′ so that $l_2' = l_1$, in which case the image, of height y_2', is in the same place as y_2. This is simply an inverted

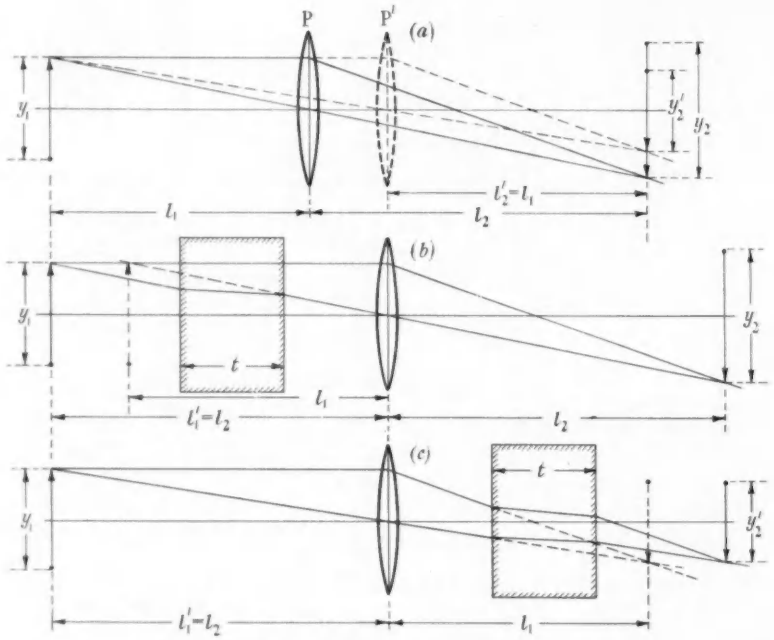


Fig. 8.—Variable magnification using a single lens.

The combined frames of the received picture are shown in fig. 7 and this should be compared with fig. 3(*a*). Although registration has been achieved the picture appears distorted, but this is due to exaggeration. In a practical case the picture tapers by only about 2·5 per cent (*i.e.*, $\varepsilon \sim 0{\cdot}025$), which is small enough to be quite unnoticeable and certainly less than deflectional errors in most reproducing apparatus. This trapezium coefficient is exceeded on the screen of most cinemas in which a tilted projector is mounted in a high projection room. Curvature of inclined boundaries is entirely negligible.

This mode of scanning was adopted in the present transmitter, but the design of optical apparatus to work at two different but constant values of magnification over consecutive periods of 0·02 sec. is an entirely

system, and the magnifications m and m' are related by the equations :—

$$\frac{y_2}{y_1} \cdot \frac{y_2'}{y_1} = mm' = 1{\cdot}0 \quad . \quad . \quad (17)$$

It is the displacement of the lens by a distance $(l_2 - l_1)$ which presents the mechanical difficulty, especially if the motion is to be discontinuous. Figs. 8(*b*) and 8(*c*) show the adopted method of virtual displacement as applied to this simple example. If $l_2 > l_1$, as is shown, then both sides are fixed at this maximum value l_2.

Clearly if the object side (or image side) is shortened to l_1 then the condition for m (or m') is satisfied. This shortening is produced by the intro-

duction of retardation blocks which must be of refractive index μ greater than the surrounding medium. The optical path reduction for a thickness t is $t(\mu - 1)/\mu$ and this must be $(l_2 - l_1)$.

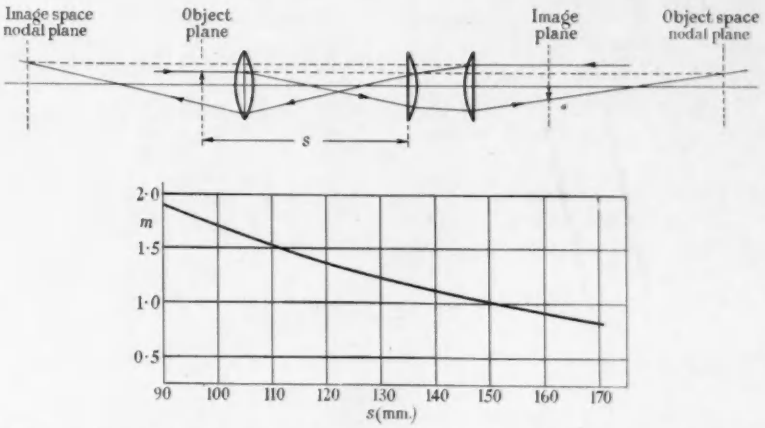


Fig. 9.—Simple constant-focus optical system with adjustable magnification.

Such a simple optical system is not sufficiently free from aberrations, and it does not provide the desirable facilities of independent controls for magnification and focus. It is possible to design a compound lens system which, by adjustment of one or more components, gives first- and second-order changes of magnification and focus respectively in the image plane, and, by the adjustment of different components, changes only the focus to a first order. Such a lens was designed and is shown in fig. 9, but the freedom from aberrations is only sufficient to meet the requirements of synchronizing channel optical systems. Ultimately a high-aperture fully corrected system[10] was obtained with the required double adjustment, and the structure can be seen as $L_2 + L_3 + L_4$ in fig. 10. The system is built up of four achromatic components of which the centre two are moved together in a barrel for magnification adjustment, and the fourth provides focus control.

(3) OPTICAL SYSTEM.

(3.1) Luminous Transfer.

Modern high-gain photocells using secondary emission electrodes can deal adequately with the smallest light flux that is ever likely to be available from a disc scanner, but a minimum limit is set to the primary current by the signal/noise ratio in the vision channel. The following considerations are necessary in a determination of the light flux to the photocell.

$$\text{The scanning hole velocity} = 2\pi rpq \quad \cdot \quad \cdot \quad (18)$$

$$\text{The scanning field width} = \frac{2\pi rq}{N} \quad \cdot \quad \cdot \quad (19)$$

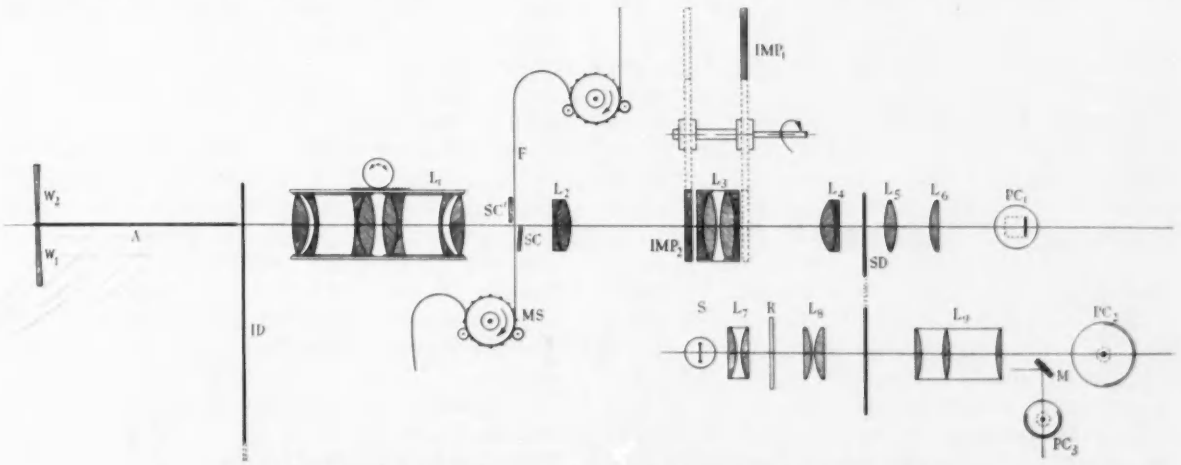


Fig. 10.—Complete optical system for intermittent magnification scanner.

in which

r = average radius of scanning holes

q = number of turns of spiral per picture

p = number of pictures per second
N = number of lines per picture

If the field aspect ratio (*i.e.*, width/height) is a, then the side and diameter of the square scanning aperture and equivalent circular hole are respectively

$$\frac{2\pi rq}{N^2 a} \text{ and } \frac{2\cdot 26\pi rq}{N^2 a} \quad \cdot \quad \cdot \quad (20)$$

$$\text{The hole area} = \frac{4\pi^2 r^2 q^2}{N^4 a^2} \quad \cdot \quad (21)$$

If the projection system has a magnification m, then the aperture area referred to the film plane is

$$\frac{4\pi^2 r^2 q^2}{N^4 a^2 m^2} \quad \cdot \quad \cdot \quad \cdot \quad (22)$$

The illumination at the film gate is

$$E = \frac{B\pi}{4K^2} \quad \cdot \quad \cdot \quad \cdot \quad (23)$$

in which B = source brightness
K = f/number of illumination system.

It is advisable to consider that the projection lens has been designed to work without cut-off, and that the aperture of the illuminant is restricted to a value K to meet this condition. In the usual mirror-condenser type of arc system the mirror images the arc crater substantially at infinity, and K is then the ratio of condenser-gate distance to effective condenser diameter.

The luminous flux through the film picture element is then

$$L = \frac{B\pi^3 r^2 q^2 \eta}{N^4 a^2 K^2 m^2} \quad \cdot \quad \cdot \quad (24)$$

η is the optical transmission efficiency from arc crater to photocell, and includes the loss due to "tunnel effect" described in sub-section (11·1). A normal light loss due to air-glass refractive index discontinuities for a single lens (2 air-glass surfaces) is about 8–10 per cent (or 0·9 db.), and as there are 34 such surfaces in the complete I.M. optical system described later, the total loss cannot be less than about 14–16 db.; in fact the actual loss is nearer 18·5 db., including the absorption of clear film in the gate. For the disc constants of sub-section (4·1) the effective scanning-hole diameter is 31 microns (μ) which is obtained approximately by a hole 37·5 μ diameter in a window 25 μ thick, owing to the apparent size reduction due to tunnel effect.

The luminous flux reaching the photocell is given by equation (24) and, using a low-intensity arc of 16,000 candles/cm.^{2}, $L = 0\cdot0095$ lumen for $K = 2$ and $\eta = 0\cdot12$ and a 37·5-μ hole. This flux is not all available if the illumination compensation device (sub-section 3·3 below) is used, as two half-brightness images must be formed on the film, or interceptor plane imaged on the film; the f/number of the illumination system is increased from $K = 2$ to about $K = 3$ so that the available "white" signal flux is $L_{max} = 0\cdot0042$ lumen. It is shown later that this luminous output is quite adequate to produce a high signal/noise ratio.

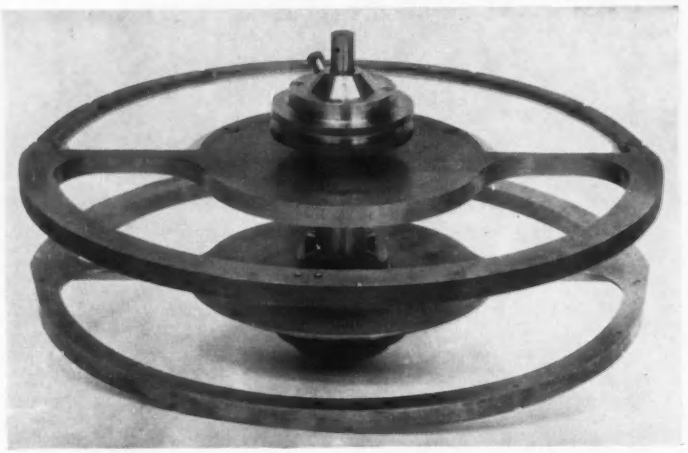


Fig. 11.—Assembly of intermittent magnification plates.

(3.2) Complete Optical System for Intermittent Magnification.

The layout of the scanner optical system is shown in fig. 10. A mirror-condenser type of low-intensity arc running at 50 amperes is used as the light source but is not included in the diagram. As the signal/noise ratio of the transmitter is very high (see sub-section 6·1) a 900-watt tungsten projector lamp is satisfactory for nearly all applications of the apparatus. The photocell PC_1 has a caesium-on-silver cathode which has a definite peak response at a wavelength of about 0·75 μ, so that at its working surface brightness a tungsten lamp produces a larger photoelectric current than an arc source attenuated to the same brightness. This is due to the lower colour temperature of the tungsten. Before the efficiency of a source to be used with a normal caesium photocell cathode can be assessed it is necessary to multiply the quoted (visual) brightness by a factor, which is largest for tungsten filaments and lowest for high-pressure mercury discharges. The factor for an arc source is between these two values.

The illumination system comprises two prisms W_1 and W_2 separated by an opaque apron A. This arrangement, together with the film-shrinkage and focus compensator SC, are described in detail in the following sub-sections.

The illuminated spiral slot of a 750-r.p.m. interceptor disc ID is imaged in the film plane by a unity magnification lens system L_1. The intermittent

magnification system proper consists of components L_2, L_3 and L_4 so mounted that L_3 is movable to provide a fine adjustment of average magnification, and L_4 is used for accurate setting of focus in the scanning disc (SD) plane. The system is so designed[10] that the distance between the object space and image space nodal planes is as great as 91·75in., while the object-image spacing is only 11·1in. The focal length of the arrangement is 24·9in., which, in conjunction with the wide spacing of the nodal planes, accounts for the freedom from focus errors due to film movement or mechanical mis-alignment of film and disc planes.

It is necessary that the projection system should be substantially free from barrel distortion. This error is proportional to the cube of the radius vector r of any point from the optical axis in the image plane. The effect is to cause a small departure from exact registration of the interlaced frames, but measurements indicate the maximum error to be not more than 0·1 per cent of the picture side. Even this error can be reduced to zero across the picture centre by adjusting the two magnification values by a small amount.

The associated intermittent magnification plates IMP_1 and IMP_2 are of thickness $t = d\mu/(\mu - 1)$, in which length d is the required virtual movement of lens combination L_3 determined on an optical bench with precision distance gauges. The refractive index μ was found to be 1·5222 at $\lambda = 0{\cdot}75\ \mu$, so that t was specified as $6{\cdot}965 \pm 0{\cdot}007$ mm. over the whole of the area. The complete assembly of two plates, of outside radius 156 mm., is shown in fig. 11.

The plates are quite thin and are situated in parts of the optical path where the angles of rays to the axis are a minimum, and for first-order imagery each plate is equivalent to an air path of t/μ. It can be shown that third-order aberrations appear which have a sign as for a negative lens. These aberrations due to the parallel plates could be corrected in the computation of the associated lenses,[11] but this has not been found necessary in the system of fig. 10.

Lenses L_5 and L_6 collect all light passing through the scanning holes and image the uniform exit pupil of L_1 on the cathode of the photocell, so that non-uniformity of sensitivity over the cathode area cannot give rise to flicker.

The film gate takes the form of a metal slide containing two equal trapezoidal apertures for the two scanning fields. When each is projected according to the appropriate magnification, registration is obtained with the required scanning field in the disc plane, and the received picture is exactly rectangular. In this way the slightly sloping edges of the I.M. distortion diagram of fig. 7 are removed.

So far it has been assumed that, in a projection system, light travels from the film to the disc, but this is not a fundamental requirement. For example, in the I.M. arrangement the disc and film could be interchanged if the optical parts are suitably adjusted.

Fig. 10 includes the complete synchronizing optical system. An exciter lamp S is imaged by L_7 near a rectangular aperture in R, which in turn is imaged by L_8 in the disc plane. The images of line and frame synchronizing holes when passing the illuminated area are produced separately on the cathodes of photocells PC_2 and PC_3 where pulse currents are originated.

(3.3) Uniformity of Illumination.

It is known that a gradual change of brightness over a picture is tolerable even if the range is as much as 4 : 1. In cinema practice, where light efficiency is very important, the positive carbon of the light source is imaged so near to the film that the screen illumination variation amounts to 2 : 1 in many cases. Even if this could be accepted for television pictures a much greater difficulty can arise when interlaced signals are required.

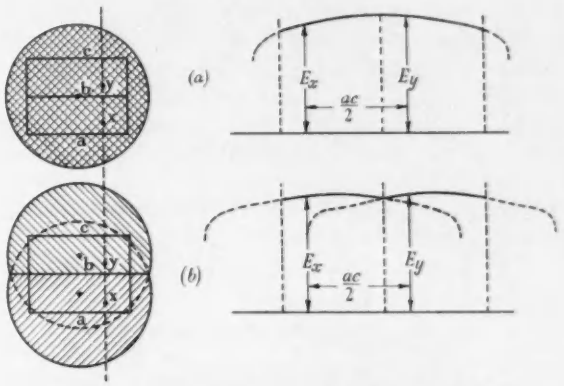


Fig. 12.—Gate illumination conditions.

A given variation becomes more apparent as the brightness change per unit length is increased, until a discontinuity is reached in the limiting case. This can occur in an interlaced picture when the brightnesses of adjacent lines, belonging to consecutive frames, are not the same. Viewed from a distance the picture seems to be constituted of half the number of lines, and a half-frame frequency flicker is introduced over the areas of non-uniformity.

In all systems using continuously moving film and a stationary source image the picture is scanned a plurality of times in different positions within the gate opening. Neighbouring points on adjacent lines of the picture must then be derived from equally illuminated points in the film plane which are separated in space by half the gate height, as one scan interval has elapsed between them and the film has moved by half a picture height in this time. This means that the gate illumination at any point in the lower scanning field must be the same as that for a corresponding point displaced from it by half

the gate height along the direction of film motion. It is accepted that some variation can occur along transverse lines through these two points if the change of illumination is the same for both lines. The best adjustment that can be obtained with most light sources is such that the image illumination has axial symmetry, but this does not meet the above conditions.

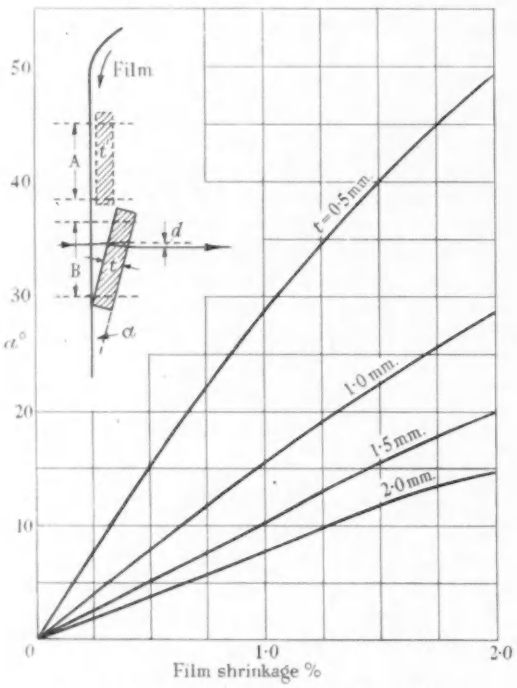


Fig. 13.—Film shrinkage compensation.

A device[12] has been developed which overcomes this difficulty, and which ensures that any brightness variations over corresponding points of the two fields are coincident. It can be seen from fig. 10 that the arc crater is imaged in the plane of the interceptor disc, and again in the gate by means of an optical system of unity magnification. The illumination non-uniformity is shown in fig. 12(*a*) for an uncorrected system. Corresponding points x and y in the two frames are subject to different illumination levels E_x and E_y. When the correction is applied the whole of the light from the arc is divided between two thin "Hysil" prisms (*ca.* 2·5 deg.) situated about 250 mm. in front of the interceptor disc. The prisms are arranged with the horizontal thin edges together, and a vertical slide is provided so that the prism assembly can be adjusted to give equal flux through each prism. The angle is chosen so that each prism produces a displacement of the image by one-quarter of the gate height as shown in fig. 12(*b*). A horizontal opaque apron passing between the prisms, and extending to the centre of the double image, ensures that no light overlaps. Each scanning field is then covered by a part of a single image and the distributions in the two fields are identical (*i.e.*, $E_x = E_y$ for all corresponding positions of x and y). Vertical movement of the prism assembly simply alters the ratio E_x/E_y to another constant value.

(3.4) Film Shrinkage and Focus Compensation.

Accuracy of interlacing depends on the maintenance of a definite relationship between film speed and the radial velocity of the disc spiral. It was found that even the most exact continuous traction film drive suffered from an inherent defect due to the variable picture height on normal film. This trouble does not occur during cinema projection as the film is moved exactly by four sprocket holes for each motion of the Maltese Cross mechanism, and the screen picture is steady and simply altered in scale by film shrinkage. In a film scanner the film fails to reach its correct position in the second scanning field by an amount directly proportional to shrinkage, as the "sprocket holes per second" and not the film speed is a constant. This is fundamental, and is not related to the co-operation of shrunk film with a perfect sprocket. In cinema practice this imperfect meshing sets a limit to the amount of shrinkage which can be tolerated. New nitrate stock is punched very accurately, but an allowable shrinkage of up to 1·5 per cent occurs with normal running and storage. One method of correction introduced an adjustment of film image velocity, in the disc plane, by alteration of the average magnification of the main optical system,[13] but in the final apparatus an optical displacement of the second scanning field was used.[14] This device, shown in fig. 13, is simply a tilting glass plate-displacer which can be adjusted by a knob on the gate. The displacement is varied until the second scanning field (on the film) is exactly 9·5 mm. away from the first, and the curves give values of tilt required to correct for any usable shrinkage.

$$d = t\left[1 - \frac{\cos\alpha}{\sqrt{(\mu^2 - \sin^2\alpha)}}\right]\sin\alpha \quad \cdot \quad (25)$$

The existence of a glass plate in part of the main optical system would introduce a focus error, but this could be corrected by another plate across the first field and by moving back the whole optical system a distance $t(\mu - 1)/\mu$. However, it was found that the use of the rotating intermittent magnification plates produced a slight focus change for the two fields, and to correct this the compensator plates were made of different thicknesses t and t', with the thicker variable in angle and in that part of the path of the optical system requiring the minimum object-image spacing. In the limiting case the thinner can be of zero thickness if the tilting

plate can give enough displacement without excessive values of α.

(4) THE SCANNING DISC.

(4.1) Specification of Disc for Intermittent-Magnification Method.

The following is an exact specification of the

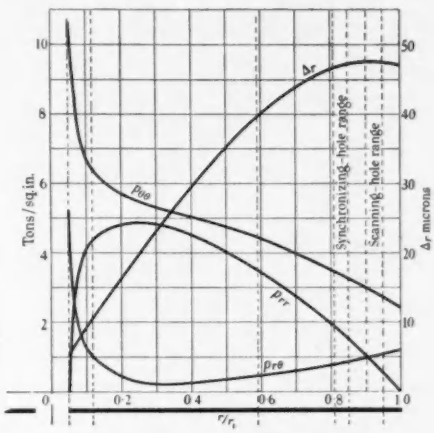


Fig. 14.—Stresses and extensions of disc of fig. 28.

I.M. disc shown in fig. 28, and used in the scanner of fig. 29.

Steel plate thickness	=	0·2 mm.
Outside radius	=	257·5 mm.
Radius of outside hole (r_1)	=	244·294 mm.
Radius of mean hole (r_{203})	=	238·000 mm.
Radius of inside hole (r_{405})	=	231·868 mm.
Radial increment δ_A	=	0·030357 mm.
Radial increment δ_B	=	0·031160 mm.
Projection magnification m_A	=	0·6473
Projection magnification m_B	=	0·6645

(4.1.1) Stress and Strain.

It can be shown[15] that at a radius r the radial stress p_{rr} and the hoop or tangential stress $p_{\theta\theta}$ are given by

$$p_{rr} = \frac{(3+\sigma)\rho\omega^2 r_2^2}{8g}\left[1 + \frac{r_1^2}{r_2^2} - \frac{r_1^2}{r^2} - \frac{r^2}{r_2^2}\right] \quad \cdot \quad (26)$$

and

$$p_{\theta\theta} = \frac{(3+\sigma)\rho\omega^2 r_2^2}{8g}\left[1 + \frac{r_1^2}{r_2^2} + \frac{r_1^2}{r^2} - \frac{r^2}{r_2^2}\left(\frac{1+3\sigma}{3+\sigma}\right)\right] \quad (27)$$

in which r_1 and r_2 are the inner and outer radii of the disc, and

σ = Poisson's ratio
ρ = density of the material
ω = angular velocity
g = gravitational acceleration.

The maximum value of p_{rr} occurs at a radius $r = \sqrt{(r_1 r_2)}$, and the function $f(rr_1r_2)$ within the square brackets is then $(1 - r_1/r_2)^2$. The maximum value of $p_{\theta\theta}$ is at the inner radius $r = r_1$ and

$$f(rr_1r_2) = 2\left[1 + \frac{r_1^2}{r_2^2}\left(\frac{1-\sigma}{3+\sigma}\right)\right] \quad \cdot \quad (28)$$

Fig. 14 shows actual values of the stresses across a diameter of the disc specified above.

The radial extension Δ_r of the disc is

$$\Delta_r = \frac{A}{8}\left[-r^3 + b_1 r + \frac{b_2}{r}\right] \quad \cdot \quad \cdot \quad (29)$$

in which $$A = \frac{(1-\sigma^2)\rho\omega^2}{Eg}$$

$$b_1 = r_2^2\left(\frac{3+\sigma}{1+\sigma}\right)\left(1 + \frac{r_1^2}{r_2^2}\right)$$

$$b_2 = r_2^4\left(\frac{3+\sigma}{1-\sigma}\right)\frac{r_1^2}{r_2^2}$$

and E is Young's modulus.

Actual values of this extension are also shown in fig. 14 and it will be seen that the maximum is not at the outer radius. At first sight this seems paradoxical, but it can be explained by considering that the centrifugal loading vanishes at the edge of the disc.

All scanning holes are situated in a region about the maximum of this curve, and the radial extension produces substantially no relative radial displacements between holes but only a shift of the scanning field as a whole.

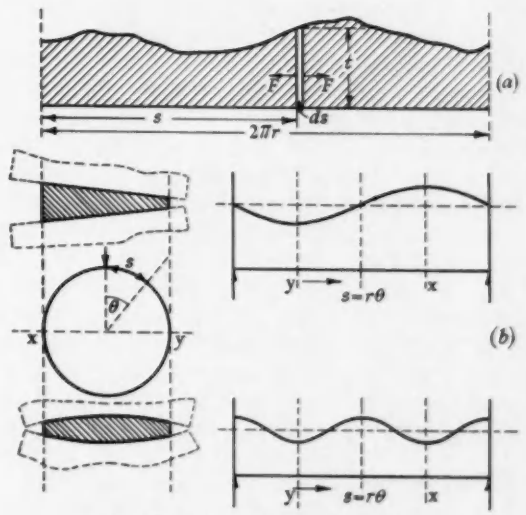


Fig. 15.—Disc thickness variation.

(4.1.2) Circumferential extension at rim.

The total circumferential strain Δ_θ at any radius r is clearly $2\pi\Delta_r$, in which Δ_r is given by equation (29). Ordinary values of this strain for a perfectly regular disc are of no consequence within the elastic

limits of the material. Many efforts have been made to ensure perfect axial symmetry of the material by such methods as criss-cross rolling, and precision lapping or grinding to constant thickness. The latter process is definitely a disadvantage as the uniformity of the surface skin of the material is disturbed and buckling is inevitable.

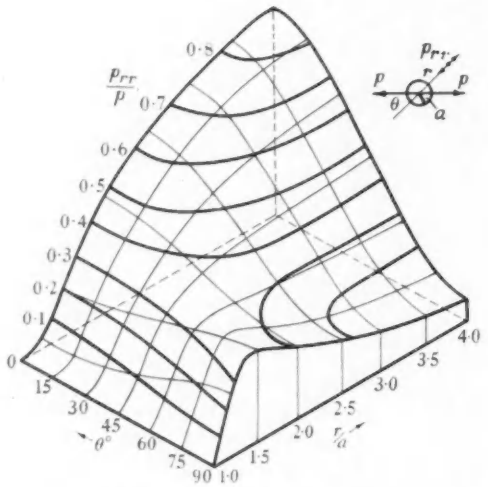


Fig. 16.—Radial stress : values of p_{rr}/p.

In nearly all cases it will be found that any departure from constant thickness of a sheet takes the form either of a very small taper from one side to another, or a cylindrical flattening towards the ends of a diameter at right angles to the line of rolling. A small cyclic variation from the average thickness has very little influence on the total circumferential extension, but the strain varies along the rim and is inversely proportional to the thickness. The extension resulting from the integration of this strain over a sector of a given angle then depends on the position of the sector. The effect is that scanning holes spaced by appreciable angles suffer relative angular displacements.

Fig. 15 shows a thickness distribution for the mean scanning-hole circle. For simplicity it is assumed that the radial width of the annular ring (normal to the diagram) is unity, and that the force F is constant along the length $2\pi r$. The extension of an infinitesimal section is $kFds/t$ in which $t = f(s)$.

The total extension is :—

$$\int_0^{2\pi r} \frac{kF}{f(s)} ds = \Delta_\theta \sim 2\pi\Delta_r \quad . \quad . \quad (30)$$

The average thickness t_0 is given by :—

$$t_0 = \frac{1}{2\pi r}\int_0^{2\pi r} f(s) \, . \, ds \quad . \quad . \quad (31)$$

and the extension of a uniform disc of thickness t_0 is

$$[\Delta_\theta]_0 = \frac{2\pi r k F}{t_0} \quad . \quad . \quad . \quad (32)$$

The disc error μ_θ is assessed as the difference between the extensions of the average and actual discs over some length $s = l$.

$$\text{Then} \quad \mu_\theta = \frac{kFl}{t_0} - \int_0^l \frac{kF}{f(s)} \, . \, ds \quad . \quad . \quad . \quad (33)$$

A disc with a simple thickness taper can be defined by $f(s) = t_0\{1 + m \sin (s/r)\}$. The error increases with l and reaches a maximum when $l = \pi r$.

$$\mu_\theta = \frac{kF}{2t_0} - \frac{kF}{2\pi r t_0}\int_0^{\pi r} \frac{ds}{1 + m \sin (s/r)} \quad . \quad . \quad (34)$$

The disc material is so chosen that the thickness variation coefficient m is small (*e.g.*, $< 0{\cdot}03$) and equation (34) reduces to

$$\mu_\theta = 2m\Delta_r \quad . \quad . \quad . \quad . \quad (35)$$

This is the maximum linear displacement of a hole with reference to a diametrically opposite hole taken as the datum.

In an actual example at 6,000 r.p.m. Δ_r reaches a maximum value of 45 μ (fig. 14), and if m is taken as $0{\cdot}03$ then $\mu_\theta = 2{\cdot}7\ \mu$, which is $7{\cdot}2$ per cent of a hole diameter. This amount is small, but it must be remembered that displacements vary as the square of the rim velocity and certain forms of disc suffer an appreciable error if the ultimate resolution, and hence speed, is attempted. The error μ_θ only

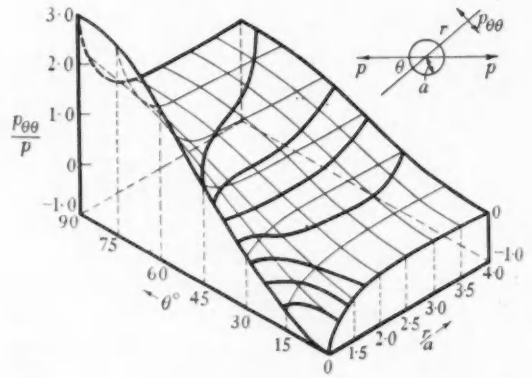


Fig. 17.—Tangential stress : values of $p_{\theta\theta}/p$.

matters if it concerns holes which produce neighbouring lines on the reproduced picture, and the trouble is avoided if the disc rotates an integral number of times per frame so that such holes are almost adjacent. When a non-integral arrangement is

used the error must be tolerated unless, by control of the form of thickness variation, μ_θ can be made to pass through minimum values when the integration is performed over the sectors between corresponding holes. For the case in which the number of turns n per picture is odd (*e.g.*, $2\frac{1}{2}$ turns per frame) the thickness function $t = t_0\{1 + m \sin (s/r)\}$ produces the error $\mu_\theta = 2m\Delta_r$, but a change to

$$t = t_0\{1 + m \sin (2s/r)\}$$

—the cylindrical taper mentioned above—gives $\mu_\theta = 0$ for the important integration interval $l = \pi r$, and a maximum $\mu_\theta = m\Delta_r$ for $l = \frac{1}{2}\pi r$.

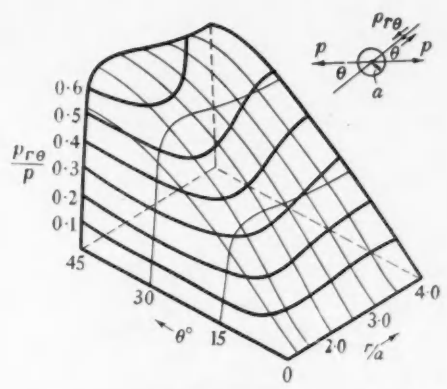


Fig. 18.—Shear stress : values of $p_{r\theta}/p$.

(4.1.3) The effect of holes.

It is not possible to construct a disc without any holes in the material, and even if the holes are small, or so spaced that there is a negligible reduction of disc cross-section, it is necessary in any strength calculations to allow for the excess stresses which appear near such holes.

The presence of the central hole for hub mounting produces a sharp rise of tangential stress near the axis, as shown in fig. 14. This stress is a maximum for the whole disc and departs very little from a value

$$p_{\theta\theta} = \frac{(3 + \sigma)\rho\omega^2 r_2^2}{4g} \quad . \quad . \quad (36)$$

for a considerable range of hole diameter. It has been shown[16] from strain considerations that there is a discontinuity in the stress in a perfectly homogeneous material when the hole vanishes, and the maximum value of $p_{\theta\theta}$ (at the centre) is reduced to $(3 + \sigma)\rho\omega^2 r_2^2/(8g)$. An infinitesimal hole at the centre gives the original result of equation (36), so it can be assumed with normal materials that the strength of a disc with no hole is somewhat less than doubled.

Holes in other parts of the disc give rise to stress concentrations and it can be shown that if a hole is punched in an infinite sheet with a principal stress p (at great distances from the hole) then at a distance r from the centre of the hole, of radius a, the stress distribution is given by :—

Radial Stress.

$$p_{rr} = \frac{p}{2}\left[\left(1 - \frac{a^2}{r^2}\right) + \cos 2\theta\left(1 - 4\frac{a^2}{r^2} + 3\frac{a^4}{r^4}\right)\right] \quad (37)$$

Tangential Stress.

$$p_{\theta\theta} = \frac{p}{2}\left[\left(1 + \frac{a^2}{r^2}\right) - \cos 2\theta\left(1 + 3\frac{a^4}{r^4}\right)\right] \quad . \quad (38)$$

Shear Stress.

$$p_{r\theta} = -\frac{p}{2}\left[\sin 2\theta\left(1 + 2\frac{a^2}{r^2} - 3\frac{a^4}{r^4}\right)\right] \quad . \quad . \quad (39)$$

in which θ is the angle between the radius vector and the principal stress. Values of p_{rr}/p, $p_{\theta\theta}/p$, and $p_{r\theta}/p$ are given in figs. 16, 17 and 18, and it can be seen that a tangential stress of three times the original stress appears at the radius of the hole.[17] Fortunately at right angles to this vector there is a stress ratio of $-1{\cdot}0$, so that if the sheet is subjected initially to two equal principal stresses the stress ratio is reduced to $2{\cdot}0$. In the general case of principal stresses p_1 and p_2 ($p_1 > p_2$) the maximum stress concentration is $3p_1 - p_2$. This is a very serious factor which probably decides the ultimate strength of most discs, and care must be taken in choosing the positions of all holes. Holes near the centre should preferably be covered by accurately machined clamping plates, and all others located in positions of equal, or small, principal stresses.

(4.1.4) Criteria for failure.

It is difficult to analyse the exact conditions which arise in a high-speed disc and result in ultimate rupture. It is generally agreed[18] that in ductile materials elastic failure takes place at a maximum value of shear stress. As both stresses $p_{\theta\theta}$ and p_{rr} in a disc are tensile, the shear stress $p_{r\theta} = \frac{1}{2}(p_{\theta\theta} - p_{rr})$ can never exceed one half of the greater. Most suitable materials for discs have an ultimate shear strength which is greater than half the tenacity, and so the maximum principal stress (usually represented by $p_{\theta\theta}$) is taken as the failure criterion in present work on discs. The presence of holes probably determines the strength against rupture, but the effect of the very high local stress concentrations is probably to introduce non-elastic strains which reduce the tendency of the material to tear from the edge of a hole. If these strains extend over distances comparable with the disc thickness, the disc, especially when stationary, will probably show buckling. This alone is sufficient to justify the practice of overspeed running of disc blanks before the construction is completed. The effect is exaggerated

in materials in which there is a badly defined yield point and a tendency to permanent set.

It is expected that the present peripheral speed of 500 ft./sec. for mild steel could be increased to at

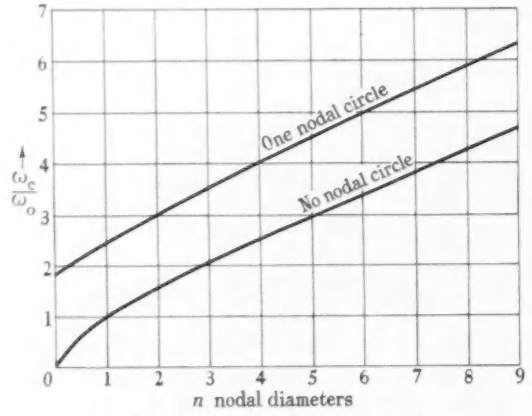


Fig. 19.—Modes of disc vibrations.

least 1,000 ft./sec. by using one of the new aluminium alloys, and such a disc would provide four times the picture elements demanded by present standards.

(4.1.5) Vibration.

Before it can be decided that a disc is usable up to a maximum speed determined only from centrifugal stress considerations it must be decided that the range is free from speeds at which vibrations can build up to a noticeable amplitude. A circular disc has a doubly infinite set of natural modes of vibration characterized by the number of nodal diameters and the number of nodal circles present. Any one of these modes can be considered as a standing wave produced by the superposition of a pair of waves travelling

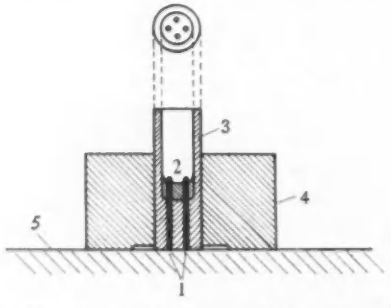


Fig. 20.—Lapping jig for tungsten punches.

1 = punch rod.
2 = hard wax.
3 = copper cylinder.
4 = polishing jig.
5 = flat plate.

round the disc in opposite directions, and if the velocity of propagation of one of these waves is equal to the peripheral velocity of the disc then the disc is liable to vibrate spontaneously under the excitation of periodic forces. Air in a disc chamber introduces viscous drag and the possibility of vortices which modify the conditions of vibration so that nodal diameters can rotate.

Centrifugal and other forces tend to stiffen the disc and raise the frequencies of vibration, while the shearing forces due to air drag tend to produce elastic instability or collapse of the disc and lower the frequencies of vibration. An analysis is available[19, 20] for the case of the disc in which the elastic rigidity is negligible and the stiffening is due only to centrifugal forces. This is certainly applicable to the scanning disc in which the ratio of diameter to thickness is as great as 2,500. The vibration frequencies ω for the disc can be calculated from the approximate formula $\omega^2 = \omega_c^2 + \omega_e^2$, in which ω_c = frequency neglecting rigidity and ω_e = frequency under elastic forces only.

The condition that $\omega_c \gg \omega_e$ is normally satisfied. General expressions for ω_c are complicated, but two important cases for disc vibration with no nodal circle and with one nodal circle are respectively :—

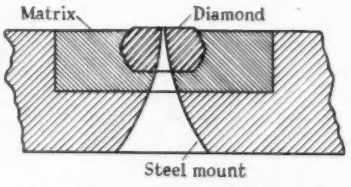


Fig. 21.—Cross-section of diamond die.

$$\frac{\omega_c^2}{\omega_0^2} = \frac{n}{4}\{(n+3) - \sigma(n-1)\} \quad \cdot \quad \cdot \quad \cdot \quad (40)$$

$$\frac{\omega_c^2}{\omega_0^2} = \frac{1}{4}\{(n^2+9n+12) - \sigma(n^2-3n+4)\} \quad (41)$$

in which ω_0 is the angular velocity of the disc and n is the number of nodal diameters. Possible conditions of oscillation arise when ω_c/ω_0 and n are integers simultaneously. The possibility of such coincidence is apparent from fig. 19.

It has been found that the scanning disc at 6,000 r.p.m. develops vibrations with slowly rotating nodal diameters when the chamber air pressure exceeds about 3 mm. of mercury, and this oscillation subsides when the pressure is reduced below 1·0 mm. The effects of temperature stresses have not been included in any calculation of strain or vibration. The disc is finished in a dull black enamel, and even in the absence of convection more than 300 watts of incident energy is required to raise the temperature by 100 degs. C.

(4.2) The Punching of Scanning Holes.

At the beginning of the development it was decided to produce the scanning holes in very thin metal diaphragms covering relatively large holes

punched in the disc material. The size of the hole required is 37·5 μ diameter and the diaphragm is of cold-rolled platinum foil, 25 μ thick and having a very fine crystalline structure. It was found that the best holes were produced by very small punches and dies.

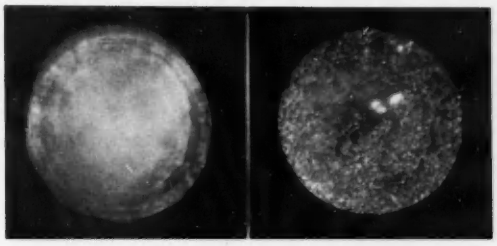


Fig. 22.—Diamond die (magnification 700).

Fig. 23.—Quartz die (magnification 700).

(4.2.1) Punches.

The material chosen for the punch[22] is specially drawn tungsten filament wire, selected for uniformity of surface finish, roundness and diameter. Diameter steps obtainable may be as small as 0·5 μ. Experiments have shown that in order to produce a 37·5-μ hole the optimum diameters of punch and die are 35 μ and 40 μ respectively. This combination gives substantially perfect shear, but the punching clearance of only 2·5 μ introduces great difficulties of alignment.

The wire is coated with electrolytic copper, in a constant-current-density plating bath, until the rod diameter is 1·0 mm. These punch rods (1) are fixed

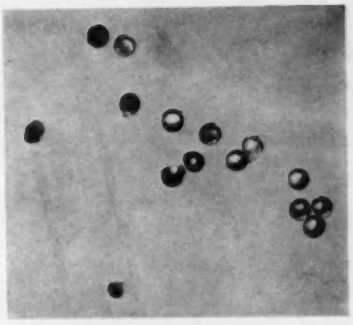


Fig. 25.—Platinum punchings, 37.5 μ diameter.

by hard wax (2) in a copper cylinder (3) as shown in fig. 20. The copper cylinder is arranged to slide freely in a polishing jig (4), and the cylinder and punch rods are lapped back with suitable abrasives against a flat plate (5) until the ends of the tungsten wire and the copper support have a highly polished optical surface.

The copper support round the tungsten must be removed for a length of about 0·1 mm. This is done by coating the side of the cylinder with wax,

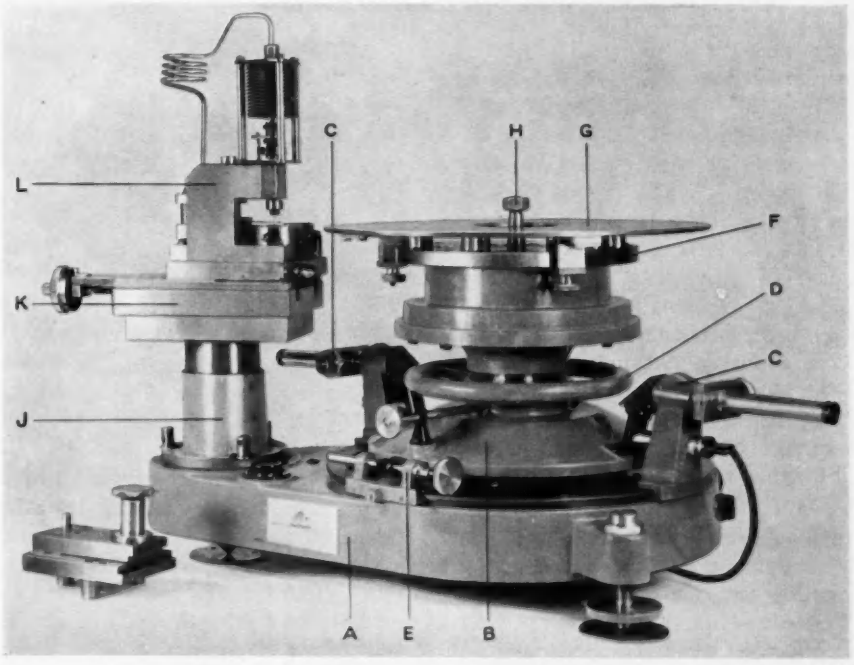


Fig. 24. The punching head and dividing engine assembly, showing also the window and the synchronising hole punch.

leaving only the polished end face clean, and immersing the whole in a bath of nitric acid agitated by compressed air.

As the platinum foil diaphragms used on the scanning disc are fixed over a window hole that is only 1·0 mm. in diameter and 0·2 mm. deep, it is necessary to taper the punch support down to 0·5 mm. to allow it to enter the window aperture.

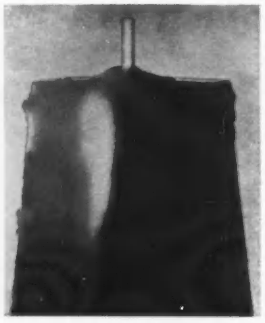


Fig. 26.—Tungsten punch, 35 μ diameter.

(4.2.2) Dies.

It has been found that the diamond is a very satisfactory material for dies of this size. The die is similar to those used in certain wire-drawing equipment and is illustrated in fig. 21. It can be seen that, by lapping back the surface of the die, the hole diameter may be adjusted to any suitable size greater than the original. Dies of this type give a sharp cutting edge and may be as small as 10 μ diameter. Fig. 22 shows a diamond die of 40 μ.

As an alternative to diamond, fused quartz or glass capillary tube may be used.[23] This is heated and worked in such a manner that the bore of the tube collapses, forming a long tapering hole down the centre. The tube is cut off and lapped back at right angles to the bore until a suitable size of hole is obtained. Care must be taken in the manipulation to ensure that the hole taper takes a definite form and is substantially free from longitudinal striations. An example of such a die is shown in fig. 23 which illustrates the precise and definite form of the circular cutting edge. The granular material (< 0·25 μ) within the hole is a special rouge which has been used in the lapping process. Dies as small as 6 μ have been made by this method.

(4.2.3) Punching head.

For the satisfactory co-operation of the punch and die the mounting and the associated moving parts must be of a very rigid construction. The main unit of the punching head is milled out of a solid block of mild steel. The punch is held in a two-jaw chuck with a parallel bore to avoid bending the punch when clamped. This chuck is fitted to the end of a round hardened-steel ram which is located near its ends by V bearings.

The punch is operated by means of oil-filled bellows situated with their line of action above that of the ram, the working force being applied between the top of the ram and the bearing block to avoid any bending of the punching head. Such an hydraulic system for producing spring-resisted mechanical movement ensures that equal and opposite forces are applied within the structure of the punching head and dividing engine assembly, so that no relative displacement occurs between the punching line and the disc axis.

The die is mounted in a hardened and polished steel block which is adjustable over a small range by means of four hardened steel set-screws, each bearing on the side of the mounting block. Location in a horizontal plane is obtained by means of spring loading against two stops, one of which is adjustable to impart a shift of the whole head so that the punch lies on the exact line of radial traverse (see subsection 4·2·5). The complete assembly can be seen in fig. 24.

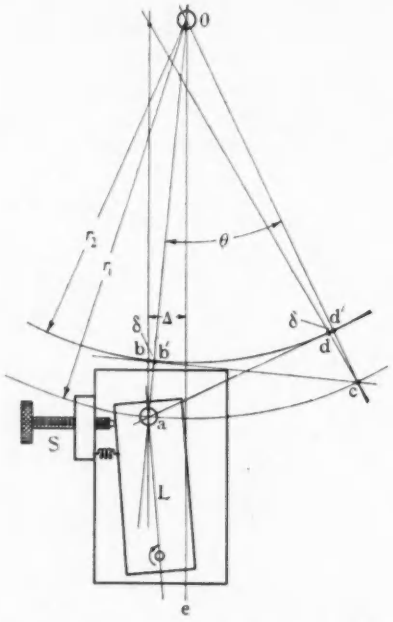


Fig. 27.—Determination of radial traverse.

In the adjustment of punch and die alignment two microscopes are set up to view the punching clearance along two axes at right angles. At the same time the depth of punch penetration is fixed by a calibrated stop.

Specimen platinum punchings 37·5 μ diameter, and from a diaphragm of 25 μ thickness, are shown in fig. 25, but the photograph is somewhat marred by meniscus films of oil at the edges of the punchings.

(4.2.4) The dividing engine.

The dividing engine used for setting the angular positions of the scanning holes is modified to suit the specialized requirements. It consists of a very rigid cast-iron bed (A) on which is mounted the vertical spindle assembly carrying the master dividing circle. A metal cover (B) over the scale is equipped with windows to enable the scales to be read by means of measuring microscopes (C) fitted to the machine bed. The spindle and circle rotate together, and this motion is adjustable by means of a hand wheel (D). Final adjustment of the angular position is obtained by clamping the cover to the main spindle and moving a micrometer screw (E) which moves the whole assembly relative to the bed.

The upper end of the spindle carries a hardened and ground steel plate (F) to which is attached a large steel disc (G) upon which the scanning disc to be punched is laid and subsequently clamped by means of a cover plate. A taper spindle (H) extends through the centre of (G) and locates the scanning-disc hub exactly in the centre of the dividing spindle.

The main scale is divided into 360 degrees and sub-divided into 4-minute divisions. Each viewing microscope has mounted in its barrel a micrometer slide which moves two vertical and parallel hair-wires across the field of vision. These hair-wires are spaced apart so that an engraved line in the main scale will just appear to lie between the wires. One complete revolution of the micrometer drum on each microscope moves the wires a distance equivalent to a main scale rotation of 120 seconds of arc. The micrometer scales are so divided that the accuracy of division obtainable is $\pm$ 1·0 second of arc, which corresponds to about $\pm$ 3 per cent of a hole diameter.

The steel pedestal (J) carries the micrometer radial slide (K) on which is mounted the punching head (L).

(4.2.5) Determination of radial traverse.

The fixing of the punching head to the micrometer radial slide is not sufficient to ensure that the punch moves along a line passing through the disc centre. Some line 0e in fig. 27 passes through the centre O and is parallel to the slide guides, and the punch must be moved to this line by rotation of the head L about an axis. The distance of the punch from the centre must also be determined and adjusted to the datum.

A disc of normal size is prepared with four platinum diaphragms, with centres approximately at positions a, b, c and d in the diagram, so disposed that pairs lie on radial lines, and lines joining diagonal pairs are tangents to the circle through the inner pair. A hole is punched at "a," the slide is moved a distance $r_1 - r_2$ and a second hole is punched at "b." The dividing head is turned by a known angle θ, and two other holes "c" and "d" punched as before. If the first two holes lie on a radial line, then so will the second pair, and measurements by means of a travelling microscope will indicate equal lengths of diagonals of the punched figure. In general this

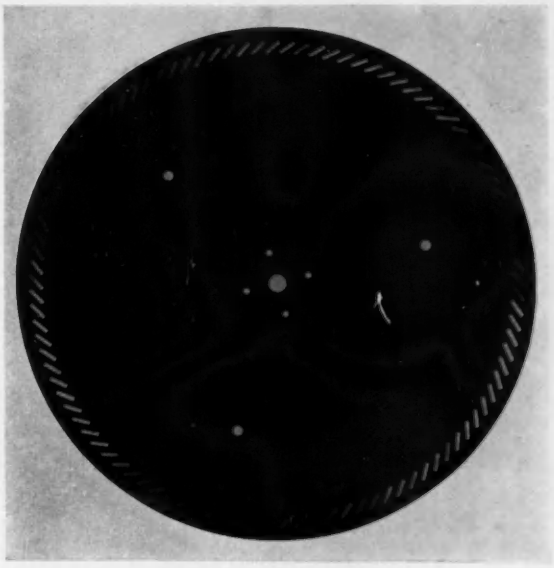


Fig. 28.—I.M.-type scanning disc for 405 lines.

equality is not realized, as both b and d are spaced from correct points b′ and d′ by an error δ. The diagonals then differ by 2δ. The punch error is given by :—

$$\Delta = \delta\left(\frac{r_1}{r_1 - r_2}\right) \quad \cdot \quad \cdot \quad (42)$$

Radii r_1 and r_2 have been given approximate values, but an exact determination is available from :—

$$r_1 = \frac{\text{length ac}}{2 \sin \frac{1}{2}\theta} \quad \cdot \quad \cdot \quad \cdot \quad (43)$$

Slide displacement $r_1 - r_2$ is of course known accurately from the micrometer readings. The required positioning of the punch is obtained by moving screw S by an amount Δ in the appropriate direction. The radial position is established, as the micrometer reading at the exact radius r_1 is known.

(4.3) Production of a Complete Disc.

The material used for the disc is best quality steel sheet, 0·2 mm. in thickness and specially selected for its uniformity and finish.

A hole is punched in the centre of a sheet and a temporary steel mandrel fitted. The hole will eventually locate the disc on its hub. Using this mandrel as a centre bearing, the disc is nibbled out of the sheet and the rough edge cleaned up by hand, after which the hub is temporarily fitted to the disc. This hub is of hardened steel with a tapered bore,

and a clamp plate is provided for holding the disc in position. It is essential that the hub itself should be perfectly true and accurately balanced so that, when the disc is added, any residual unbalance is due entirely to the disc. The disc is now mounted on the dividing engine and clamped into position by means of a cover plate. A small manually-operated punch is mounted on the radial slide and adjusted for radial traverse in the manner already described for the punching head. The 4-turn spiral of 1·0 mm. diameter picture windows is punched according to the angle and radius figures for the scanning holes.

Platinum foil 25 μ thick is cut into strips for covering four windows each.[24] The foil diaphragm is fastened to the disc by soft solder, and it is necessary that the cleaned disc should be tinned. During this operation, and the subsequent one of sweating the diaphragms on to the disc, it is desirable that no high-temperature gradients, which could lead to permanent distortions on cooling, should be established in the disc material. This is avoided by making use of what is in effect an isothermal process which consists of placing the disc on the surface of a large electrically heated hot-plate, and gradually bringing the temperature up to the running temperature of solder. Powdered solder and flux are wiped round the disc in a uniform film of very small thickness. The hot-plate is afterwards reduced to a temperature just below the melting point of the solder film, and the platinum strips sweated over the groups of picture windows by means of a hot soldering bit of aluminium. By this means it is possible to reduce the local temperature differences to as little as 15–20 deg. C.

The three concentric circles of 1·0 mm. diameter line-synchronizing holes, and the short length of frame-synchronizing holes, are produced by the window punch.

The datum angular position of all the holes punched in the disc is not random but is chosen for reasons of balance. All hole arrays are unsymmetrical and the resultant out-of-balance effect has been assessed. The disc blank is measured for unbalance and the datum fixed so that the unbalance forces are opposite in sign. In this way the residual unbalance is reduced to a minimum. The whole of the disc, except for the platinum strips, is given a hard-drying matt black finish. The disc is finally balanced during the finishing process, and any small adjustment necessary is made by uniformly increasing the enamel thickness on both surfaces of the lighter part of the disc. The disc is then ready for the punching of the scanning holes.

The whole of the punching operation is done in a temperature-controlled room in order to avoid differential expansions of all parts of the complete apparatus shown in fig. 24.

Special precautions must be taken against dust, which can cause irregularities of punching or blocking of the scanning holes. It is assumed that partial blocking of only one hole is sufficient to render the disc unusable. Numerous methods of cleaning, and maintaining the disc clean, have been tried with varying degrees of success. A good method is to draw a volatile solvent through the punched hole by means of a small suction chest mounted on the stage of a microscope so that the whole operation can be watched until complete. A much more efficient and rapid method is under development, and results give promise of a final solution of the dust problem.

When housed in a scanning-disc chamber a disc has remained perfectly clean for as long as 2 years, and no signs of deterioration are apparent. A finished I.M. disc is shown in fig. 28.

(*To be continued*).